

VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS
INFORMATIKOS KATEDRA

**Konvoliuciniai neuroniniai tinklai smegenų navikų
segmentavimui magnetinio rezonanso tomografo
nuotraukose**

**Brain tumor segmentation of MRI images using convolutional
neural networks**

Magistro baigiamasis darbas

Atliko:	Rokas Gipiškis
Darbo vadovė:	Prof. dr. Olga Kurasova
Recenzentas:	Dr. Rimantas Kybartas

Vilnius – 2020

Santrauka

Darbe nagrinėjamos pažangiausios konvoliucinių neuroninių tinklų architektūros ir jų pritaikymas smegenų navikų segmentavimui magnetinio rezonanso tomografo nuotraukose. Identifikuojami modeliai pastaraisiais metais pasiekę geriausius Dice'o panašumo koeficiento (F_1) rezultatus semantinio segmentavimo užduotyse. Taip pat eksperimentiškai tiriami 3 neuroninių tinklų architektūriniai sprendimai (*U-Net*, *U-Net* su dėmesio vartais ir *SegNet*) siekiant pagerinti žemo lygio gliomų segmentavimo rezultatus smegenų aksialinių pjūvių nuotraukose. *SegNet* tinklo segmentavimo rezultatai pagerinami tinklą modifikavus pridėdant rekurentinius liekanų blokus.

Raktiniai žodžiai: konvoliuciniai neuroniniai tinklai, magnetinio rezonanso tomografija, semantinis segmentavimas, *U-Net*, *SegNet*

Summary

State-of-the-art convolutional neural network architectures and their application to brain tumor segmentation of MRI images are analyzed. We identify models that over the last few years have achieved the best results in semantic segmentation tasks in terms of Sorensen-Dice coefficient (F_1). We also experiment with 3 neural network architectures (U-Net, Attention U-Net and SegNet) in order to improve results of the segmentation of low grade gliomas in the axial plane. Segmentation results of SegNet are improved by modifying the network with recurrent residual blocks.

Keywords: convolutional neural networks, magnetic resonance imaging, semantic segmentation, U-Net, SegNet

Turinys

1. Įvadas	5
1.1 Darbo tikslas	8
1.2 Darbo uždaviniai	8
2. Literatūros analizė	9
2.1 <i>U-Net</i> tinklo architektūra: enkoderis ir dekoderis	10
2.1.1 Praktiniai <i>U-Net</i> taikymai gliomų segmentavime	11
2.2 <i>Segnet</i> tinklas	12
2.3 Kiti sprendimai enkoderio-dekoderio tinkluose	14
2.3.1 Dvimatė enkoderio-dekoderio architektūra	14
2.3.2 Trimatė architektūra	14
2.3.3 Trimatė architektūra su erdviniu sutelkimo sluoksniu	15
2.3.4 <i>VGG</i> ir dviejų kelių <i>DeepMedic</i> tinklų įtakoti trimačiai sprendimai	19
2.4 Architektūrų palyginimas	22
3. Eksperimentiniai tyrimai	24
3.1 Duomenų rinkiniai	24
3.2 Duomenų paruošimas	25
3.3 Įvertinimo kriterijai	25
3.4 Modeliai	27
3.4.1 <i>U-Net</i> tyrimas ir rezultatai	28
3.4.2 <i>U-Net</i> su dėmesio vartais tyrimas ir rezultatai	32
3.4.3 <i>SegNet</i> tyrimas ir rezultatai	34
3.4.4 <i>SegNet</i> su rekurentiniais liekanų blokais tyrimas ir rezultatai	35
3.5 Rezultatų įvertinimas ir palyginimas	37
4. Išvados	39
Literatūra	40
Priedai	44

1. Įvadas

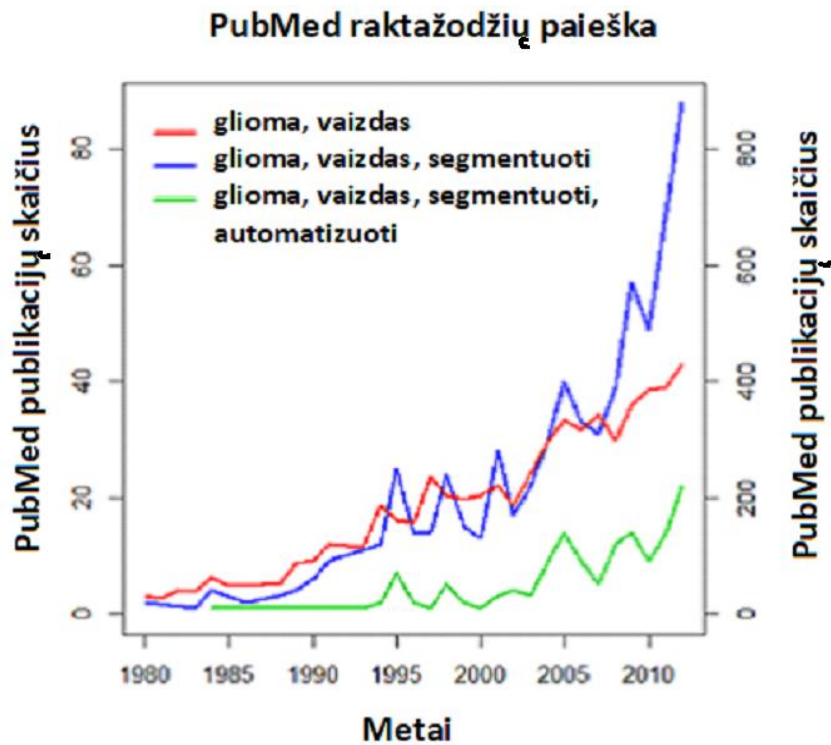
Šiame darbe apžvelgiamas konvoliucinių neuroninių tinklų (KNT) taikymas smegenų navikų magnetinio rezonanso tomografo (MRT) nuotraukų analizei, koncentruojantis į semantinį segmentavimą, kuris yra esminis biomedicininis vaizdų analizėje. Didžiausias dėmesys skiriamas smegenų navikų gliomų segmentavimo problematikai. Dėl netaisyklingos gliomų formos ir didelės jas sudarančių sričių įvairovės šis uždavinys yra laikomas vienu sudėtingiausių medicininių vaizdų analizėje [CVG+18]. Gliomos yra dominuojanti piktybinių smegenų auglių rūšis, sudaranti 80 proc. jų atvejų bei apie 30 proc. visų centrinės nervų sistemos ir smegenų navikų [GJ12]. Jos skirstomos į žemo lygio gliomas (ŽLG) ir aukšto lygio gliomas, kurios apibūdina patį agresyviausią smegenų naviko tipą. Nepaisant paprastesnės vidinės ŽLG struktūros, lyginant su aukšto lygio gliomomis, ŽLG segmentavimo užduotis yra sudėtingesnė dėl mažesnio navikų dydžio ir mažo kontrasto [LWY+17].

Susidomėjimas automatinio segmentavimo pritaikymu smegenų navikų analizei vis auga (1 pav.). Kaip pastebima [MJB+15], su gliomų segmentavimu susijusių publikacijų skaičiaus augimas jau daugiau kaip dešimtmetį yra didesnis nei būtų tiesiniu atveju. Smegenų navikų segmentavimo uždavinio sprendimo metodai įprastai skirstomi į dvi pagrindines grupes [CPA+16]:

1. Generatyvinius, besiremiančius ankstesnėmis žiniomis apie skirtingų audinių tipų išvaizdą ir jų pasiskirstymą.
2. Diskriminatyvinius, tiesiogiai išmokstančius santykį tarp vaizdo bruožų ir segmentavimo žymenų.

Būtent pastarųjų – diskriminatyvinių modelių – kategorijai priskiriami sprendimai, pasitelkiantys konvoliucinius tinklus vis tikslesniam gliomų ir jų sričių segmentavimui atlikti. Kompiuterinės regos srityje dirbančių tyrėjų vienareikšmiškai pripažįstama, jog šiandien KNT yra efektyviausia priemonė segmentuojant medicininius vaizdus. Šių tinklų raida įgavo pagreitį nuo 2012 m., kuomet buvo pristatytas Alexo Krizhevsky'io, Ilya'os Sutskeverio bei Geoffrey'io Hintono [KSH12] pasiūlytas *AlexNet* tinklas, *ImageNet* klasifikavimo konkurse laimėjęs pirmą vietą ir ženkliai pagerinęs tų metų vaizdų atpažinimo rezultatus. KNT tyrimų srityje tinklų architektūrų plėtros mastai išties dideli – *AlexNet* buvo sudarytas iš 8 sluoksnių, o jau po kelerių metų pristatyta *ResNet* sudarė 152 sluoksniai [HZR+15].

Magnetinio rezonanso tomografai yra dažniausiai naudojamas įrenginys smegenų analizei atlikti. Priešingai nei kompiuterinė tomografija, jo naudojimas nedidina rizikos sveikatai. MRT vaizdai naudojami identifikuojant tokius smegenų veiklos sutrikimus kaip Alzheimeris, epilepsija, šizofrenija ir išsėtinė sklerozė.



1 pav. Su glioma ir vaizdų segmentavimu susijusių raktažodžių, įvestų *PubMed* paieškos variklyje iki 2012 m. skaičius. Šaltinis: [MJB+15].

Semantinio segmentavimo užduotis reikalauja, jog išvesties rezultatas būtų tankus (angl. *dense*). Tankumas pasireiškia tuo, jog KNT išvestyje turime gauti vaizdą, kurio kiekvienas pikselis būtų priskirtas tam tikrai klasei (t.y. įvyksta kiekvieno pikselio klasifikavimas). Tuo tarpu lengvesnėje vaizdo klasifikavimo užduotyje išvestyje gautume ne vaizdą, o žymenį, nurodantį prognozuojamą įvesties klasę arba klasių tikimybes, gaunamas pritaikius *softmax* arba *sigmoid* funkciją. Priešingai nei klasifikavimui skirti tinklai, segmentavimui pritaikytos architektūros turi ne tik enkoderio, bet ir dekoderio dalį, kuri yra reikalinga siekiant išlaikyti pradines nuotraukų dimensijas. Todėl, kaip pastebi Zahangiras Alomas su bendraautoriais [AHY+18], lyginant su klasifikavimui skirtais tinklais, segmentavimo modeliai turi beveik dvigubai daugiau tinklo parametrų.

Nepaisant plataus MRT naudojimo, lieka svarbių, bet neišspręstų problemų, siekiant automatizuoti MRT vaizdų apdorojimą ir analizę. Pasak šioje srityje dirbančių tyrėjų [AGH+17], „automatinė smegenų struktūrų segmentacija ir anomalijų aptikimas išlieka neišspręsta problema dėl įprastų anatominių variacijų smegenų morfologijoje, variacijų MRT skaitytuvuose, vaizdų gavimo trūkumų ir variacijų patologijos formoje“. Dabartinėje praktikoje MRT vaizdų analizė visų pirma yra atliekama gydymo įstaigų ekspertų. Tai daug laiko, žinių ir pasirengimo reikalaujantis procesas. Tačiau ir tuomet vyrauja subjektyvumo veiksnys, todėl keletą medikų išvados gali skirtis. Tuo tarpu didesnės MRT vaizdų analizės automatizavimo galimybės, pasitelkiant konvoliucinius tinklus, leistų už MRT vaizdų analizę atsakingam personalui efektyviau priimti sprendimus.

Tyrinėjant MRT nuotraukas susiduriama su mažesnėmis prieinamomis duomenų aibėmis (pavyzdžiui, lyginant su *ImageNet* konkurse pateikiamais duomenimis). Iš dalies tai galima paaiškinti siekiu užtikrinti su asmens sveikata susijusių duomenų privatumą. Tačiau pastaruoju metu vis daugėja bandymų apjungti viešai prieinamų MRT vaizdų duomenų archyvus ir leisti objektyviau išmatuoti skirtingų jų analizei taikomų metodų tikslumą, ir tai patvirtina vis augantį susidomėjimą dirbtinių neuroninių tinklų pritaikomumu MRT vaizdų analizei. Kaip tokių bandymų pavyzdžius galima išskirti 2013 m. Japonijoje organizuotą *MRBrains*¹ projektą, apjungiantį į automatizuotą MRT vaizdų segmentaciją besigilinančius tyrėjus ir leidžiantį jiems pasiūlyti savo sprendimus *MRBrainS* pateiktai MRT vaizdų duomenų aibei ir palyginti jų gautų segmentacijos rezultatų tikslumą, bei nuo 2012 m. vykdomą *BraTS*² (angl. *Brain Tumor Segmentation*) smegenų navikų segmentavimo konkursą.

Dėl palyginti naujo (arba – tiksliau – naujai atrasto) susidomėjimo konvoliuciniais tinklais, jų pritaikymo MRT vaizdų analizei galimybių tyrimas yra dar ganėtinai naujas. Iš išsamesnių pastarųjų metų apžvalgų, besigilinančių į skirtingų KNT architektūrų panaudojimą smegenų struktūrų analizei MRT, galėtume išskirti 2019 m. tyrimą [LL19] bei 2018 m. *Big Data Mining and Analytics* žurnale pristatytą apžvalgą [LPL+18]. Taip pat paminėtinas ir 2017 m. *Journal of Digital Imaging* publikuotas tyrimas [AGH+17] apie dirbtinių neuroninių tinklų panaudojimą MRT vaizdų segmentacijai kiekybinėje MRT analizėje. Visi trys tyrimai daro palankias išvadas dėl tolimesnės giliojo mokymo pritaikymo MRT smegenų nuotraukų analizei raidos.

¹ <https://mrbrains13.isi.uu.nl/>

² <http://braintumorsegmentation.org/>

1.1 Darbo tikslas

Nustatyti konvoliucinių neuroninių tinklų tipus, labiausiai tinkamus smegenų navikų segmentavimui magnetinio rezonanso tomografo nuotraukose, ir pasiūlyti tinklo modifikaciją, leisiančią tiksliau segmentuoti žemo lygio gliomas.

1.2 Darbo uždaviniai

1. Atlikti literatūros, susijusios su konvoliucinių neuroninių tinklų architektūromis, naudojamomis smegenų navikų MRT nuotraukų segmentavime, analizę.
2. Identifikuoti konvoliucinių neuroninių tinklų tipus, labiausiai tinkamus MRT smegenų navikų nuotraukų segmentavimui bei nustatyti jų ypatumus.
3. Programiškai realizuoti kelias perspektyviausias architektūras, iš kurių tikimasi geriausių rezultatų tikslaus vaizdų segmentavimo prasme.
4. Naudojant sukurtą programinę įrangą eksperimentiškai ištirti realizuotus neuroninius tinklus, juos palyginant pagal Sorenseno-Dice'o panašumo koeficientą.
5. Pasiūlyti ir ištirti tinklo modifikaciją su labiausiai žemo lygio gliomų segmentavimui tinkamais tinklo parametrais.

2. Literatūros analizė

Šioje dalyje apžvelgta dalis mokslinės literatūros, aktualios konvoliucinių neuroninių tinklų pritaikymo magnetinio rezonanso tomografo nuotraukų segmentavimui. Kaip pastebi [LL18] autoriai, reikšminga biomedicininė nuotraukų segmentavimo tinklų dalis siūloma smegenų segmentavimo uždaviniams. Dauguma čia aptariamų sprendimų pasirinkti dėl jų autorių dalyvavimo paskutiniuose *BraTS* konkursuose, organizuojamuose Pensilvanijos universiteto medicinos institute. Šie konkursai yra žymiausi multimodalinio smegenų navikų segmentavimo srityje ir koncentruojasi į gliomų – pirminio tipo (t.y. atsirandančių smegenyse iš smegenų ląstelių, o ne dėl metastazių) navikų, dažniausiai pasitaikančių ir pavojingiausių tarp piktybinių smegenų auglių – ir jas sudarančių sričių segmentavimą bei pačių pažangiausių (angl. *state-of-the-art*) metodų šioje srityje vertinimą.

Dėl konkurso duomenų aibėje naudojamų skirtingų modalumų, dauguma čia analizuojamų straipsnių pasitelkia keturių MRT modalumų arba sekų (angl. *sequences*) – *T1*, *T2*, *T1CE* ir *FLAIR* – nuotraukas. Skirtinguose modalumuose galima matyti skirtingas naviką sudarančias sritis, todėl jų įvairovė prisideda prie segmentavimo tikslumo. Neuroninių tinklų sprendimuose modalumai paprastai traktuojami kaip spalvų kanalai. Aptariant įvesties dimensijas taip pat paminėtina, kad trimačio „nesupjaustyto“ įvesties vaizdo dydis yra 240x240x155. Segmentavimo kokybei įvertinti šis konkursas įprastai naudoja dvi metrikas: Sorenseno-Dice'o panašumo koeficientą (statistinėje binarinio klasifikavimo analizėje dar žinomą kaip F_1) bei Hausdorffo atstumą. 2018 m. konkurse prie segmentavimo vertinimo metrikų pridėtos dar ir jautrumo bei specifiškumo metrikos. 2020 m. *BraTS* konkursas tik prasideda, todėl jame naudojamos dirbtinių neuroninių tinklų architektūros, jų nustatymai bei rezultatai turėtų paaiškėti iki šių metų pabaigos, tuo tarpu detalesni sprendimai – tik 2021 m. Taigi preliminarių rezultatų su validavimo duomenų aibe turėtume sulaukti iki šių metų rudens. Verta pastebėti, jog visi paskutinių metų laimėtojų sprendimai rėmėsi KNT.

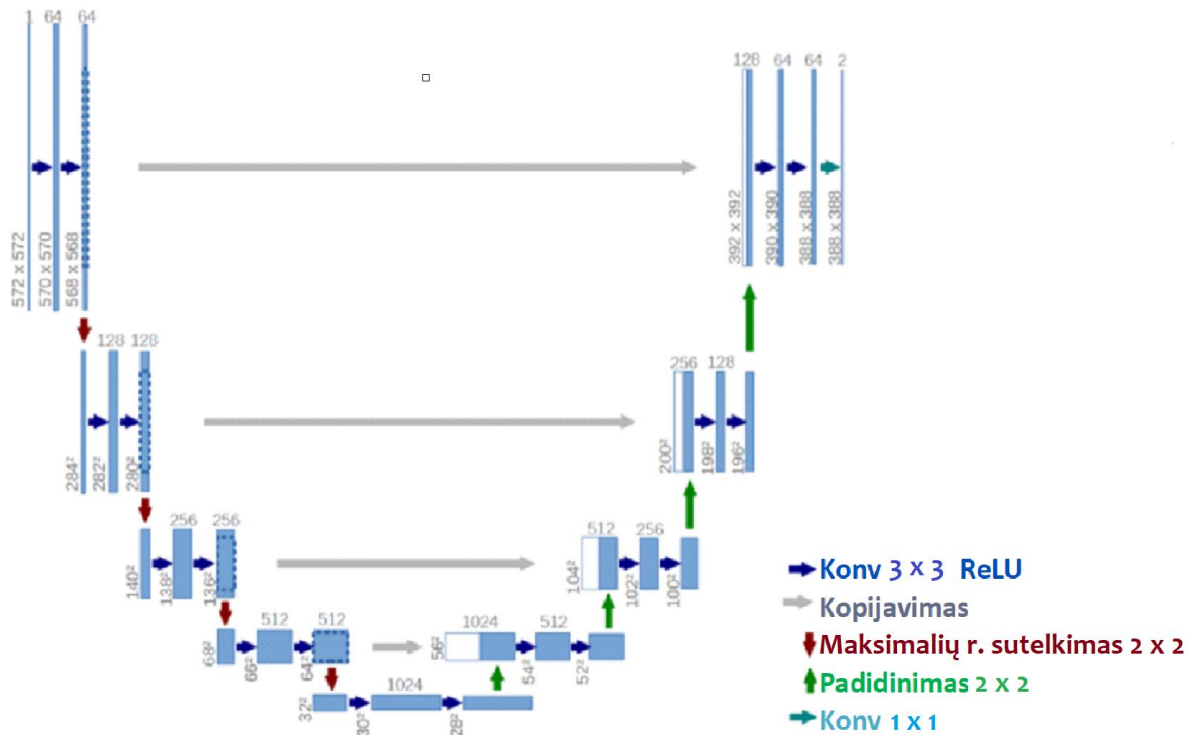
BraTS pateikia didelę profesionalių radiologų anototą ir segmentotą duomenų aibę – 2018 m. ją sudarė 210 aukšto lygio gliomų ir 75 žemo lygio gliomų atvejai. Nors šie skaičiai gali pasirodyti nedideli vertinant vaizdų analizės uždavinius bendrai, navikų segmentavimo srityje tokia duomenų aibė laikytina didele (pastarąjį apibūdinimą vartoja ir Fabianas Isensee'us su savo komanda [IKW+18]). Tai galėtų būti viena iš priežasčių, kodėl viena populiariausių KNT architektūrų smegenų segmentavimo uždaviniuose yra *U-Net* – šiuo tinklu gaunami palyginti

tikslūs segmentavimo rezultatai net ir naudojant iš nedidelio duomenų skaičiaus susidedančią duomenų aibę.

2.1 *U-Net* tinklo architektūra: enkoderis ir dekoderis

U-Net tinklas [RFB15], pristatytas Olafo Ronneberger'io ir jo bendraautorių, yra pilnai konvoliucinis tinklas (angl. *fully convolutional network*), susidedantis iš dviejų kelių, dar vadinamų enkoderiu ir dekoderiu. Kalbant apie enkoderio-dekoderio tinklą, enkoderį ir dekoderį galime suprasti ir kaip atskirus tinklus. Pirmasis kelias – enkoderis – susideda iš konvoliucinių ir maksimalių reikšmių sutelkimo (angl. *max-pooling*) sluoksnių. Šią dalį (2 pav.) vaizduoja žemyn besileidžiantys sluoksniai kairėje *U-Net* pusėje: du vienas po kito einantys konvoliuciniai sluoksniai bei po jų sekantis 2×2 maksimalių reikšmių sutelkimo sluoksnis ir dar keletas jų pasikartojimų, kol mūsų analizuojamas vaizdas tampa vis mažesnis. Tokiu būdu gaunama daugiau informacijos apie tai, kas yra vaizde (t.y. išmokstami tam tikri požymiai), bet mažiau žinoma apie tai – kur tai yra.

Dešinioji *U-Net* pusė, simetriška kairės pusės sluoksniams, vadinama dekoderiu. Ši dalis atsakinga už tikslios vietos nustatymą, kas ir yra viena svarbiausių segmentavimo užduočių. Tam reikalingas vaizdo išdidinimas, kuriam pasiekti pasitelkiamos ne tik konvoliucijos, bet ir dekonvoliucijos, dar kitaip vadinamos perkeltomis konvoliucijomis (angl. *transposed convolutions*). Jas kiek supaprastintai galima įsivaizduoti kaip atvirkštines operacijas konvoliucijai, kurios metu įvesties matrica yra sumažinama. Pavyzdžiui, jeigu taikydami konvoliucijos operaciją 4×4 dydžio matricai pritaikysime 3×3 dydžio filtrą su 1 poslinkiu (angl. *stride*), gausime 2×2 dydžio matricą. Tuo tarpu perkeltos konvoliucijos sluoksnis bandytų atgauti pradines matricos reikšmes, todėl enkoderio dalyje sumažintas vaizdas dekoderio dalyje pamažu būtų išdidintas. Nors [RFB15] siūlomas tinklas ir nėra pirmasis, pasitelkiantis enkoderį ir dekoderį, tačiau, kaip pastebi [CYA+19], *U-Net* kūrėjai juos apjungė su praleidimo jungtimis (angl. *skip connections*), kurios leidžia lengviau atgauti erdvinę informaciją dekoderyje. Šios tarp sluoksnių jungtys pažymėtos pilka spalva (2 pav.) ir veda iš enkoderio sluoksnių į atitinkamo lygmens dekoderio sluoksnius.



2 pav. *U-Net* tinklo architektūra su enkoderiu – kairėje, o dekoderiu – dešinėje pusėje.

Šaltinis: [RFB15].

2.1.1 Praktiniai *U-Net* taikymai gliomų segmentavime

Fabianus Isensee'us [IKW+18], kartu su savo bendraautoriais, pristato pilnai sujungto KNT architektūra besiremiantį *U-Net* segmentavimo sprendimą, gavusį antros vietos įvertį 2018 m. *BraTS* segmentavimo kategorijoje. Isensee'us pastebi, jog didelė dalis tyrėjų, dirbančių navikų segmentavimo srityje, bando pasiekti geresnį įvertį pasitelkdami tinklo architektūrinės modifikacijas, užuot didesnę dėmesį skyrę tinklo apmokymo etapui. Be to, pasak šio autoriaus, itin sunku pasiūlyti kažką geresnio nei *U-Net* tinklas segmentavimui atlikti. Ypač jeigu *U-Net* tinklas yra tikrai gerai apmokytas. Be to, didelė svarba skiriama duomenų normalizavimo žingsniui: „Kiekvienas kiekvieno paciento modalumas normalizuojamas atskirai, atimant vidurkį ir padalinant iš standartinio smegenų srities nuokrypio. Už smegenų esanti sritis prilyginama 0“ [IKW+18].

U-Net enkoderio dalis surenka semantinę informaciją, dėl to prarandama erdvinė vaizdo informacija, kuri, pasitelkiant praleidimo jungtis, yra atgaunama dekoderio dalyje. Taip pat Isensee'aus ir jo kolegų siūlomas sprendimas, panašiai kaip ir jo praeitų metų sprendimas, tinklo apmokyme naudoja atsitiktinai parinktas dideles įvesties vaizdo dalis (angl. *patches*) –

128x128x128. Tačiau tuomet susiduriama su trūkumu – trimačiame *U-Net* didelis vaizdo dalies dydis gali apkrauti atmintį dėl didesnės apimties skaičiavimų. Vienas iš siūlomų sprendimų – sumažinti požymių žemėlapių (angl. *feature maps*) skaičių prieš pat vaizdo išdidinimą (angl. *upsampling*). Kaip aktyvacijos funkcijos pasirenkami pralaidūs ReLU (angl. *leaky ReLU*).

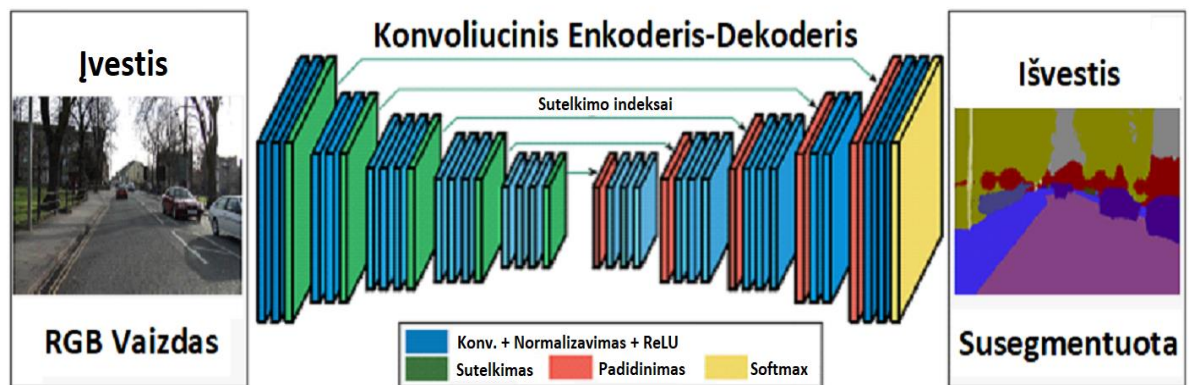
Kalbant smegenų navikų duomenų analizę po jų apdorojimo (angl. *postprocessing*), verta išskirti šią Isensee'aus pastabą: „Viena iš sudėtingiausių *BraTS* konkurso dalių yra atskirti mažas kraujagysles naviko branduolio srityje (kuri turi būti įvardyta arba kaip edema, arba kaip nekrozė) nuo naviko sustiprinimo srities. Kadangi tai ypač pavojinga žemo lygio gliomų pacientams, kurie gali ir neturėti naviko sustiprinimo, mes pakeičiam visus naviko sustiprinimo vokselius su nekroze, jei pacientas turi mažiau kaip 500 naviko sustiprinimo vokselių“ [IKW+18]. Taigi Isensee'us nusistato tam tikrą slenkstinę reikšmę, kad sumažintų su kraujagyslių atskyrimu susijusią riziką. Paminėtina, kad apmokymo metu buvo panaudota penkių sluoksnių kryžminė validacija. Apibendrinant galima teigti, kad geriausi rezultatai buvo gauti kartu apmokant (angl. *cotraining*) tinklą ir su papildomais Isensee'aus komandos duomenimis – tiesa, formaliai kiek skirtingais nuo *BraTS* pateiktų – gautais iš jų institucijų.

2.2 *SegNet* tinklas

Yi-Ju Changas su savo bendraautoriais pristato sprendimą [CLY+18], kuriame trimačius MRT nuotraukų duomenis siūloma semantiškai segmentuoti pasitelkiant dvimatį KNT *SegNet*. Šiuo atveju vaizdų augmentavimo tikslais naudojami atsitiktiniai apmokymo vaizdų kontrastų, kampų ir šviesumo pokyčiai. Šiam sprendimui tai leido pasiekti aukštesnę Dice'o įvertį lyginant su rezultatais, kai buvo naudojami vien tik *BraTS* organizatorių pateikti duomenys.

SegNet architektūra, pirmąsyk pasiūlyta Kembridžo universiteto mokslininkų Vijayo Badrinarayanan, Alexo Kendallo ir Roberto Cipolla'os [BKC16] 2015 m., yra sudaryta iš enkoderio ir dekoderio tinklų, kurių paskirtį jau apžvelgėme ankstesniame skirsnyje, nagrinėdami *U-Net* architektūrą. *SegNet* tinklo enkoderį sudaro 13 konvoliucinių sluoksnių (3 pav.), kurie sutampa su *VGG-16* tinklo, skirto objektų klasifikavimo uždaviniui spręsti, pirmaisiais konvoliuciniais sluoksniais. Tai leidžia apmokymo procesą pradėti su svoriais, apmokytais klasifikuoti dideles duomenų aibes. Dekoderio tinklas taip pat susideda iš 13 sluoksnių, o „paskutinė jo išvestis įeina į daugiaklasį *softmax* klasifikatorių, kad pateiktų klasių tikimybes atskirai kiekvienam pikseliui“ [BKC16].

SegNet kūrėjai pastebi, kad jų architektūra reikalauja mažesnio atminties saugojimo, dėl to galimas nedidelis tinklo tikslumo sumažėjimas. Tačiau jis neturėtų turėti reikšmingos įtakos praktiniam tinklo taikymui (įdomu pastebėti, kad pradinė tinklo paskirtis buvo kelio vaizdų analizė). Lyginant su kitu panašios architektūros tinklu – *U-Net*, *SegNet* naudoja maksimalių reikšmių sutelkimo indeksus, o ne perduoda pačius požymių žemėlapius atitinkamiems dekoderio sluoksniams. Nenaudodamas šių indeksų *U-Net* tinklas reikalauja didesnio atminties suvartojimo. Be to, kaip jau minėta, *SegNet* pasitelkia *VGG* tinklo konvoliucinių sluoksnių svorius.



3 pav. Segmentavimo tinkle *SegNet* architektūra. Šaltinis: [BKC16].

Toliau analizuodami savo ir kitų segmentavimo tinklų architektūrinius skirtumus, *SegNet* kūrėjai pastebi, jog didelė dalis enkoderio-dekoderio tipo tinklų turi vienodą enkoderio dalį, o esminis jų formų skirtumas glūdi dekoderio dalyje. Be to, jie palygina pilnai konvoliucinio tinklo ir *SegNet* struktūrą, pasitelkdami supaprastintas abiejų architektūrų versijas, sudarytas iš mažesnio kiekio sluoksnių. Palyginime naudojamas *SegNet-Basic* tinklas susideda iš 4 enkoderių ir tokio paties dekoderių skaičiaus. Tuo tarpu *FCN-Basic* pilnai konvoliucinis tinklas turi tą pačią enkoderio tinklo struktūrą, tačiau tinklo dekoderiai turi pilnai konvoliuciniam tinklui būdingus bruožus. Lyginant su *FCN-Basic*, privalumas pasireiškia tuo, kad *SegNet* suvartoja 11 kartų mažiau atminties. Taip yra todėl, kad *FCN-Basic* tinklas saugo pilnus enkoderio požymių žemėlapius, užuot saugojęs vien tik maksimalių reikšmių sutelkimo indeksus.

Kita vertus, autoriai pastebi, jog „*SegNet-Basic* turi dekoderį su 64 požymių žemėlapiais kiekviename dekoderio sluoksnyje. Palyginimui, *FCN-Basic*, naudojantis matmenų mažinimą, turi mažiau (11) požymių žemėlapių kiekviename dekoderio sluoksnyje. Tai sumažina konvoliucijų skaičių dekoderio tinkle ir todėl *FCN-Basic* yra greitesnis darydamas išvadas” [BKC16]. Tačiau, įskaitant *SegNet-Basic* dekoderio tinklą, ši architektūra yra sudėtingesnė už *FCN-Basic*, o tai

suteikia daugiau lankstumo ir leidžia pasiekti didesnę apmokymo tikslumą per tą patį iteracijų skaičių.

SegNet autoriai yra pasiūlę ir tinklo modifikaciją – Bajesiškąją *SegNet* [KBC16], į pradinį tinklą įtraukiantį Monte Karlo metodą. Tokios architektūros modelio tikimybinis segmentavimas leistų atsižvelgti į netikrumo laipsnį dėl prognozuojamo rezultato. Tai ypač aktualu dirbant su mažesnės apimties duomenų aibėmis. Siūlomame modelyje didžiausi netikrumo laipsniai pasireikštų ties objektų briaunomis ir ties sunkiai atpažįstamais objektais.

2.3 Kiti sprendimai enkoderio-dekoderio tinkluose

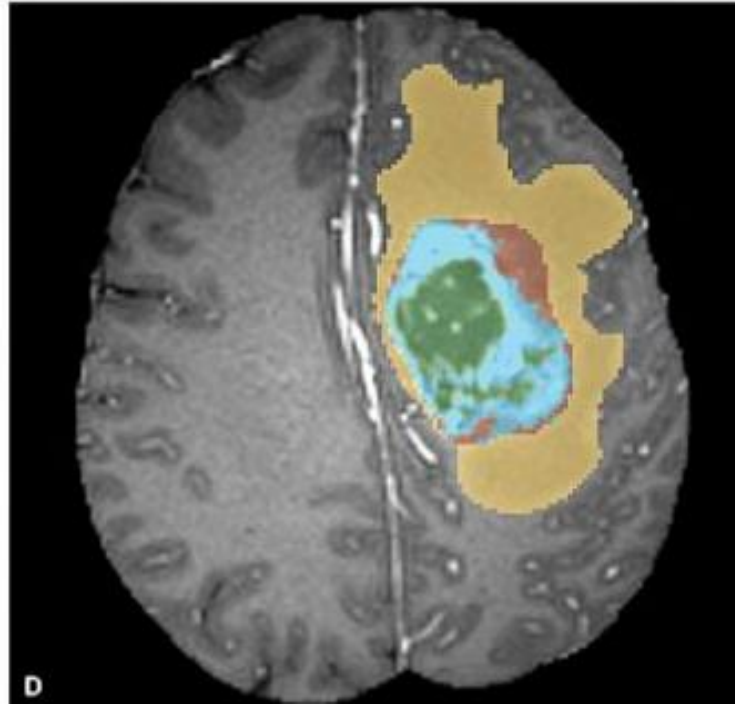
2.3.1 Dvimatė enkoderio-dekoderio architektūra

Eze Bensonas kartu su bendraautoriais [BPF+18] siūlo pritaikytą smėlio laikrodžio (angl. *adapted hourglass*) tinklą, kaip *U-Net* pagerinimą, pridėdami konvoliucijas prie praleidimo jungčių. Tai taip pat enkoderio-dekoderio tipo tinklas. Šiuo atveju tyrėjai kiekvieno MRT modalumo arba sekos duomenis normalizavo atskirai. Jie pasirinko naudoti atskirus dvimačius pjūvius, o ne trimačius duomenis, todėl sprendžiama dvimačio segmentavimo problema. Tokiu būdu atliekant skaičiavimus mažiau apkraunama atmintis. Vietoje 285 trimačių objektų, panaudojama 44175 dvimačių vaizdų aibė. Duomenų augmentacijai buvo panaudoti du metodai: vertikalus apvertimas ir atsitiktinė intensyvumo variacija. Daroma išvada, kad ir dvimatės architektūros gali pasiekti gerus rezultatus, tik jos bus sąlyginai gilesnės ir reikalaus daugiau papildomo darbo.

2.3.2 Trimatė architektūra

Andriy'us Myronenko [Myr18] pristato trimačio segmentavimo tinklą, gavusį geriausią įvertį (pagal vidutinį Sorenseno-Dice'o panašumo koeficientą visam navikui – 0,91, naviko branduoliui – 0,8668, o sustiprintam branduoliui – 0,8233) segmentavimo srityje iš 390 2018 m. *BraTS* užsiregistravusių komandų. Jis pastebi, kad kaip ir dauguma semantinio segmentavimo tinklų, jo siūlomas tinklas remiasi enkoderio-dekoderio architektūra. Myronenko aprašomoje KNT architektūroje, enkoderio tinklo dalis, atsakinga už požymių išgavimą iš vaizdų, yra didesnė lyginant su dekoderio dalimi. Taigi, kitaip nei *U-Net* tinklo atveju, enkoderio ir dekoderio dalys nėra simetriškos. Myronenko siūlomo tinklo enkoderio dalis pasitelkia *ResNet* blokus. Visų konvoliucijų dydis vienodas – $3 \times 3 \times 3$, o pradinis filtrų kiekis – 32. Kaip rezultatą enkoderio pabaigoje gauname 8 kartus mažesnę nei įvesties vaizdą su 256 požymiais. Dekoderio tinklo dalis

susideda iš padidinimų (angl. *upsizing*) ir po jų einančių blokų. Kiekvieno padidinimo metu požymių skaičius sumažinamas dvigubai, pasitelkiant $1 \times 1 \times 1$ konvoliucijas. Dekoderio pabaigoje gauname tokio pačio dydžio vaizdą, kaip ir įvestyje. Tai itin svarbu segmentuojant vaizdą. Myronenko savo tinklą paruošė naudodamas *Tensorflow* ir apmokė jį su *NVIDIA Tesla V100 GPU*.



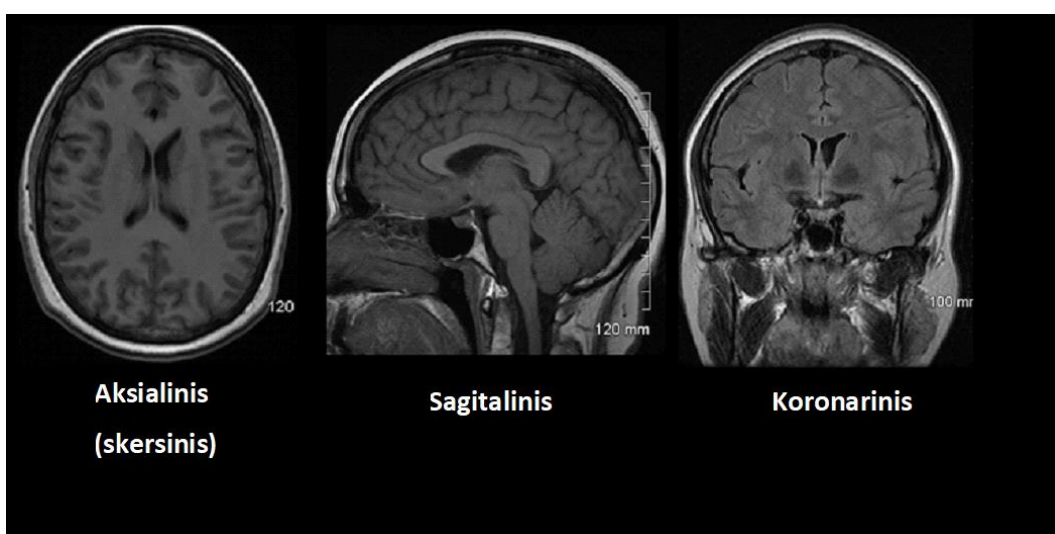
4 pav. Susegmentuota naviko sustiprinimo sritis pažymėta mėlyna spalva. Šaltinis: [MJB+15].

Tuo tarpu Mariano Cabezasas ir jo bendraautoriai [CVG+18] pristato segmentavimo metodą, sudarytą iš *U-Net* tipo trimačio tinklo su $3 \times 3 \times 3$ dydžio filtrais (angl. *kernels*) bei keturių skirtingų KNT architektūrų darinio (angl. *ensemble*). Taip pat buvo pasitelktas ir perkeliamas mokymas (angl. *transfer learning*), kuriam įgyvendint naudojamas *VGG-16* tinklas, prieš tai jau apmokytas su vaizdais iš *ImageNet* konkurso duomenų aibės. Darbe daroma prielaida, kad sunkiausia segmentuoti auglio sritis yra naviko sustiprinimo sritis (angl. *the enhancing tumor area*) (4 pav. pažymėta mėlyna).

2.3.3 Trimatė architektūra su erdviu sutelkimo sluoksniu

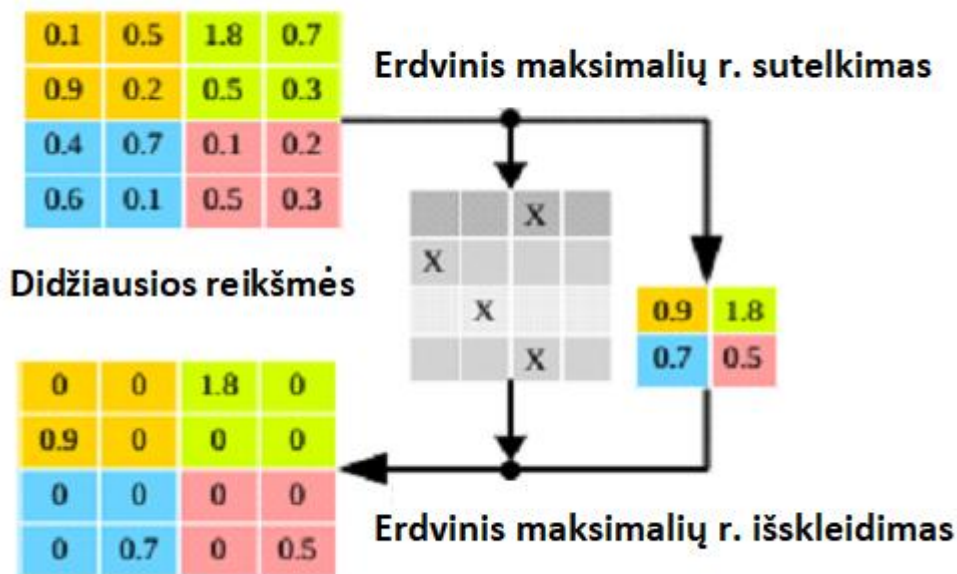
MRT vaizdus galime žiūrėti skirtingų anatominių plokštumų (angl. *anatomical planes*) pjūviuose: aksialiniame (x-z ašys), koronariniame (y-x ašys) ir sagitaliniame (y-z ašys) (5 pav.). Kadangi skirtingose anatominėse plokštumose galime geriau pamatyti skirtingas smegenų navikų

sritis, į tai galima atsižvelgti ir planuojant tinklo architektūrą. Pavyzdžiui, Subhashis Banerjee'is su bendraautoriais [BMS18] siūlo segmentavimo modelį, kuris yra sudarytas iš trijų konvoliucinių tinklų. Tokiu būdu navikas segmentuojamas trijose skirtingose plokštumose. Galiausiai trys KNT yra apjungiami, kad būtų galima gauti trijų matmenų segmentuotą naviką su jį sudarančiomis sritimis. Pasak tyrėjų, „pastebima, jog integruota prognozė iš skirtingų plokštumų yra pranašesnė tikslumo ir sprendimo tvarumo atžvilgiu, lyginant su bet kuria viena plokštuma besiremiančią prognoze“ [BMS18].



5 pav. Anatominių plokštumų pjūviai galvos smegenyse: aksialinis, sagitalinis ir koronarinis. Šaltinis: <https://sites.google.com/site/postgraduatetraining/image-acquisition/the-basics?tmpl=%2Fsystem%2Fapp%2Ftemplates%2Fprint%2F&showPrintDialog=1>.

Straipsnyje [BMS18] siūlomas modelis yra enkoderio-dekoderio tipo KNT. Enkoderio dalis įprastai pasitelkia sutelkimo sluoksnius, kad sumažintų (angl. *downsample*) vaizdą į aukštesnio lygio požymius. Sumažinimo metu dingsta erdvinė informacija, kuri medicininių vaizdų segmentavime yra svarbi vaizdo išdidinimo metu. Dėl šios priežasties autoriai naudoja erdvinį sutelkimą (angl. *spatial-pooling*), į siūlomą tinklą įvesdami erdvinį maksimalių reikšmių sutelkimo sluoksnį. Šis sluoksnis išsaugo ne tik maksimalias reikšmes, kaip yra įprasta naudojant maksimalų reikšmių sutelkimą, bet ir tų reikšmių vietas (6 pav.).



6 pav. Erdvinio maksimalių reikšmių sutelkimo operacija viršuje. Išimenuojamos didžiausios reikšmės bei tų reikšmių vietos. Šaltinis: [BMS18].

Pažymėtina, jog dažniausiai medicininių vaizdų segmentavimo uždaviniuose naudojami *U-Net* tipo tinklai. Taip pat gliomų segmentavime dažniau susiduriame su trimatėmis, o ne dvimatėmis sprendimais, nors pradinė *U-Net* architektūra [RFB15] visų pirma buvo pritaikyta dvimatėms vaizdams. Dalis *BraTS* 2018 m. segmentavimo užduotyje naudotų *U-Net* sprendimų duodavo neblogus rezultatus vienoje naviko srityje, tačiau buvo patiriama sunkumų kitose. Tokių tinklų pavyzdžiai galėtų būti Erico Caverio ir bendraautorių [CCL+18] bei Naomi Fridman [Fri18] siūlomi sprendimai. Straipsnyje [CCL+18] siūloma segmentavimui pasitelkti tris skirtingus *U-Net* tinklus – t.y. po tinklą atskirai naviko sričiai. Iš visų trijų naviko sričių, jų modelis tiksliausiai nustatė viso naviko buvimo vietą. Tuo tarpu Naomi Fridman pastebėjo, jog jos siūlomas dvimatis *U-Net* tipo sprendimas tiksliau segmentuoja didesnius nei mažesnius navikus. Kadangi įprastai navikų dydis stipriai skiriasi, Fridman svarsto galimybę naudoti atskirus specializuotus *U-Net* tinklus mažesnių navikų segmentavimui atlikti. Be to, Fridman siūlomas *U-Net* sprendimas taip pat susiduria su naviko sustiprinimo srities segmentavimo problema.

Dar vienas segmentavimo sprendimas, pasiūlytas Siddhartha'os Chandra'os su bendraautoriais [CVF+18], pristato modifikuotą trimatės *ResNet* architektūros tinklą, labiau pritaikytą segmentavimo uždavimams. *ResNet* tinklai, pasižymintys gebėjimu „peršokti“ vieną ar daugiau savo tinklo sluoksnių, dažnai naudojami su kompiuterine rega (angl. *computer vision*) susijusiose užduotyse, tačiau, pasak autorių, standartinis trimatis *ResNet* tinklas nėra pats

tinkamiausias navigu segmentavimo uždavimui spręsti, todėl yra siūlomi tinklo patobulinimai. Trimačių konvoliucijų skaičiavimai užtrunka ilgai, nes jų optimizavimo tyrimas yra tik ankstyvoje stadijoje. Autoriai bando sumažinti tinklo gylį ir padidinti jo plotį, kas turėtų padidinti tinklo tikslumą bei sumažinti jo apmokymo laiką. Dėl to naudojamas mažiausias *ResNet* tinklas – *ResNet-18*. Tirdami eksperimentiškai, *ResNet-18* davė geresnius rezultatus lyginant su gilesnėmis jų versijomis: *ResNet-34* ir *ResNet-50*.

Segmentavimo lentelėje (1 lentelė), besiremiančioje Dice'o panašumo koeficientu, matome žaliai pažymėtus pirmos ir antros vietos laimėtojus segmentavimo kategorijoje, o taip pat mažėjimo tvarka pagal geriausią viso navigu Dice'o įvertį pateiktus sprendimus. Pirmos vietos metodas parodo geresnius rezultatus segmentuojant visas tris navigu sritis.

1 lentelė. Rezultatai pagal segmentuojamą navigu sritį naudojant vidutinį Sorenseno-Dice'o panašumo koeficientą (tik straisnių, dalyvavusių *BraTS*, jeigu tokie duomenys yra prieinami).

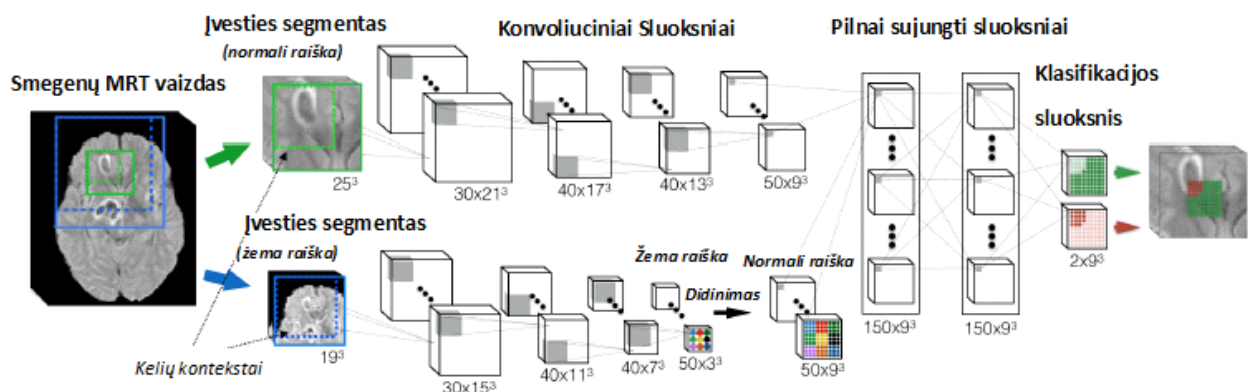
Pirmasis autorius	Visas navigas	Navigu branduolys	Sustiprintas branduolys
Myronenko	0,910	0,8668	0,8233
Isensee'us	0,9083	0,8544	0,8101
Choudhury	0,9059	0,8016	0,8154
Zhou	0,90	0,83	0,79
Cabezasas	0,8892	0,7264	0,7403
Ma	0,881	0,773	0,743
Caveris	0,878	0,760	0,724
Chandra	0,8718	0,7989	0,7405
Fangas	0,856	0,726	0,722
Changas	0,824	0,896	0,766
Tuanas	0,82	0,70	0,68
Bensonas	0,8163	0,6347	0,5910

2.3.4 VGG ir dviejų kelių *DeepMedic* tinklų įtakoti trimačiai sprendimai

Casamitjana'os su bendraautoriais [CPA+16] pristatytas sprendimas taiko trimačius konvoliucinius tinklus ir eksperimentuoja su skirtingomis pilnai konvoliucinėmis architektūromis. Kadangi trimačiai tinklai turi itin didelį parametrų skaičių, norint apmokyti modelį neviršijant turimų resursų, buvo pasitelktas mažesnis filtrų ir jų dimensijų skaičius lyginant su dvimate tinklo versija. Iš viso palyginimui pasirinktos trys trimatės KNT architektūros, kurios yra pilnai sujungtos bei geba daryti tankias išvadas (angl. *dense-inference*), t.y. daryti spėjimus apie visą trimatį darinį vienu metu.

Pirmas siūlomas modelis – *3DNet_1* – remiasi *VGG* (angl. *Visual Geometry Group*) tinklo architektūra kartu su praleidimo jungtimis. Įprastai *VGG* konvoliucinis tinklas [SZ15] sudaromas iš 16 (*VGG-16*) arba 19 (*VGG-19*) sluoksnių. Antrasis tiriamas modelis – *3DNet_2* – taip pat remiasi *VGG*, susitraukiančią jo dalį apjungdamas su daugiasluoksniu dekonvoliuciniu tinklu (išsiplečiančia dalimi). Susitraukiančios dalies sluoksniai yra sujungti su išsiplečiančios dalies sluoksniais.

Trečiojo nagrinėjamo modelio – *3DNet_3* – architektūra paremta *DeepMedic* [KLN+17] tinklu – 11 sluoksnių trimačiu KNT. Būtent šis tinklas autoriams leidžia pasiekti geriausių rezultatų – 99,71 tikslumą, naudojant 2015 m. *BraTS* duomenų aibę. Šio tinklo architektūra susideda iš dviejų skirtingų konvoliucinių kelių (angl. *pathways*) – (7 pav.). Antrasis siūlomo trimačio KNT kelias pasitelkia sumažintus (angl. *down-sampled*) vaizdus. Dvigubo kelio pasirinkimą *DeepMedic* kūrėjai paaiškina tuo, kad siekiama panaudoti tiek platesnio, tiek ir siauresnio konteksto informaciją. Platesnio konteksto informacijai būtų galima priskirti naviko sričių vietą smegenyse ir su ja susijusius požymius – jie būtų išmokstami antrame kelyje. Tuo tarpu su siauresnio konteksto informacija – pavyzdžiui, vietinių struktūrų išvaizda – susiję požymiai – pirmame [KLN+17]. Tokiu būdu siūlomas *DeepMedic* tinklas gali tuo pat metu apdoroti įvesties vaizdą skirtingose skalėse.



7 pav. *DeepMedic* kūrėjų siūlomas dviejų kelių KNT, leidžiantis tuo pačiu metu išmokti lokalius požymius viename kelyje, o platesnio masto požymius - kitame. Toks modelis yra pasitelkiamas ir Casamitjana'os komandos. Šaltinis: [KLN+17].

2017 m. Stanfordo universiteto komanda - Liyue Shen ir Timothy Andersonas [SA17] - pristatė tris naujas KNT architektūras gliomų segmentavimui bei ištyrė perkeliama mokymo (angl. *transfer learning*) galimybę, pritaikant su *BraTS* duomenų aibę apmokytus modelius *Rembrandt*³ (angl. *Repository of Molecular Brain Neoplasia Data*) duomenų aibei. [SA17] daro išvadą, jog šioje srityje perkeliama mokymo pritaikymas yra daug žadantis, tačiau vis dar reikalaujantis tolesnių tyrimų bei jų rezultatų pagrindimo. Perkeliamas mokymas itin aktualus dirbant su nesužymėtomis arba tik dalinai sužymėtomis duomenų aibėmis, tačiau medicininių vaizdų srityje jo taikymui iškyla didesnių sunkumų - pavyzdžiui, medicininių vaizdų duomenų bazės labai skiriasi priklausomai nuo pacientų ir užduočių. Taip pat autoriai tiria metodus, padėsiančius išvengti per didelio tinklo pritaikymo prie apmokymo duomenų (angl. *overfitting*), kuris gali atsirasti ir dėl nepakankamai didelio įrašų skaičiaus *BraTS* ir *Rembrandt* duomenų aibėse.

Stanfordo tyrėjai pastebi, jog lig šiol gliomų segmentavime vyrauja pikseliais besiremiantys *U-Net* tinklo taikymai: „Smegenų navikų segmentavimui naudojant daugiamodalinį MRT, buvo pasiūlyta keletas KNT besiremiančių metodų, įskaitant tuos, kurie remiasi pavienių MRT pjūvių segmentavimu, tūriniu segmentavimu bei KNT apjungtais kartu su kitais statistiniais metodais. Beveik visos dabartinės architektūros smegenų navikų segmentavimui naudoja pikseliais besiremiančiu (angl. *pixel-wise*) *U-Net* taikymu [...], kuris lig šiol buvo daug žadantis, bet vis dar rodo ribotus pasiekimus" [SA17]. Pikseliais besiremiantis segmentavimas paima nedidelę vaizdo dalį ir klasifikuoja centrinę tos dalies pikselį. Tačiau straipsnio autoriai dėmesį nusprendžia skirti didesnėmis vaizdo dalimis besiremiančiomis KNT architektūroms. Pirmieji du siūlomi KNT turėtų

³ <https://wiki.cancerimagingarchive.net/display/Public/REMBRANDT>

atlikti binarinę klasifikaciją į auglio ir ne auglio sritis. Trečiasis modelis yra pilno vaizdo pilnai KNT su penkiais konvoliuciniais sluoksniais bei 2x2 maksimalių reikšmių sutelkimu siekiant sumažinti vaizdą. Straipsnyje prieinama išvados, jog pilno vaizdo pilnai konvoliucinio tinklo rezultatai yra prasti lyginant su pirmų dviejų modelių, besiremiančių vaizdo dalimis (angl. *patch-wise*), rezultatais. Visa tai galima paaiškinti tuo, jog šio tinklo rezultatams didelę įtaką daro vaizdo fono (angl. *background*) klasės.

2 lentelė. Siūlomi KNT bei jų modifikacijos.

Publikacija	Siūlymas
Changas et al [CLY+18]	Dvimatis KNT <i>SegNet</i>
Myronenko [Myr18]	Nesimetriška enkoderio-dekoderio architektūra su <i>ResNet</i> blokais
Isensee'us et al [IKW+18]	3D <i>U-Net</i> su minimaliomis modifikacijomis
Banerjee'is et al [BMS18]	Modelis su 3 KNT su erdviniu maksimalių reikšmių sutelkimu
Bensonas et al [BPF+18]	Smėlio laikrodžio tinklas kaip <i>U-Net</i> pagerinimas
Cabezasas et al [CVG+18]	<i>U-Net</i> tipo trimatis tinklas ir skirtingų KNT darinys
Caveris et al [CCL+18]	Trys skirtingi <i>U-Net</i> tinklai – po vieną skirtingai naviko sričiai
Fridman [Fri18]	Dvimatis <i>U-Net</i> sprendimas
Chandra et al [CVF+18]	Modifikuotas trimatis <i>ResNet</i> tinklas pritaikytas segmentavimui
Casamitjana et al [CPA+16]	Trys trimačiai KNT, įtakoti <i>VGG</i> ir <i>DeepMedic</i> tinklų
Shenas ir Andersonas [SA17]	Trys KNT su perduodamu mokymu

2.4 Architektūrų palyginimas

Didžioji dalis MRT semantinio segmentavimo sprendimų remiasi enkoderio-dekoderio architektūromis, kurių ryškiausi pavyzdžiai galėtų būti medicininė vaizdų analizėje itin populiarūs *U-Net*, pilnai KNT ir *SegNet*. Kaip pastebi Isensee'us [IKW+18], visi 2017 m. *BraTS* laimėtojų sprendimai bent jau dalinai rėmėsi enkoderio-dekoderio tipo neuroniniais tinklais. Tą pačią tendenciją matome ir naujesniuose sprendimuose. Tiriant tinklų formą reikia paminėti, jog tokių architektūrų enkoderio dalis dažnai būna panaši – pavyzdžiui, *SegNet* kūrėjai pastebi, kad daugeliui segmentavimo architektūrų yra bendras tas pats enkoderio tinklas ir jos skiriasi vien tik dekoderio tinklo forma, tačiau šis skirtumas bendrai tinklo architektūrai suteikia asimetriškumo [BKC16]. Taip pat įdomu pastebėti, jog didesnė dalis mūsų analizuojamų KNT sprendimų yra trimačiai, nors gliomų segmentavimo užduočiai galimi ir dvimačiai sprendimai. Šiuos būdus galima apjungti, visų pirma analizuojant dvimačius skirtingų anatominių plokštumų pjūvių vaizdus su skirtingais KNT, o po to juos sujungiant į trimatį segmentuotą smegenų vaizdą.

Lyginant su daugiau atminties reikalaujančiu *U-Net* tinklu, *SegNet* architektūra yra mažesnė ir naudojami mažiau atminties. Tačiau tai atsiliepiama sąlyginai mažesniu tikslumu. Tuo tarpu, palyginus supaprastintas *SegNet* ir pilnai konvoliucinio tinklo architektūrų versijas matyti, jog *SegNet* yra pranašesnis, jeigu daroma prielaida, jog priimant sprendimą atmintis yra apribota ir dėl šio mažesnio atminties vartojimo yra leidžiama paaukoti dalį sprendimo priėmimo laiko. Visgi iš *BraTS* segmentavimo užduočiai pasiūlytų sprendimų matyti, jog didesnė tyrėjų dalis kaip pagrindą remiasi *U-Net* architektūra bei jos modifikacijomis.

Normalizavimas yra itin svarbus MRT duomenims, nes vaizdai dažnai skiriasi priklausomai nuo institucijos, joje naudojamų protokolų bei įrangos tipo. Apmokant didelius tinklus su nedideliu duomenų skaičiumi, reikia imtis priemonių prieš per didelį tinklo prisitaikymą prie apmokymo duomenų. Tam naudojamos duomenų padidėjimo (angl. *augmentation*) priemonės: atsitiktiniai vaizdo pasukimai, padidinimai ar sumažinimai, elastinė deformacija (ją akcentuoja ir patys *U-Net* tinklo kūrėjai) ir pan. Iš aptartų straipsnių matyti, jog gliomų segmentavimo atveju, sunkiausiai segmentuojama yra naviko sustiprinimo sritis (angl. *enhancing tumor*).

Tolesniems eksperimentiniams tyrimams kaip pagrindas pasirinkti dveji enkoderio-dekoderio architektūros KNT: *U-Net* ir *SegNet*. *U-Net* tinklas pasirinktas dėl tikslių segmentavimo rezultatų *BraTS* konkursuose (pavyzdžiui, viso naviko segmentavimo užduotyje Isensee'aus komandos siūlomo *U-Net* sprendimo Dice'o įverčio skirtumas lyginant su pirmos vietos sprendimu yra tik 0,0017) ir dėl efektyvaus prisitaikymo dirbant su nedidelės apimties duomenų

aibėmis. Tuo tarpu *SegNet* – dėl mažesnių atminties resursų reikalavimų bei mažesnio parametru skaičiaus, kuris leistų greičiau apmokyti ir ištirti skirtingas modelių versijas.

3. Eksperimentiniai tyrimai

3.1 Duomenų rinkiniai

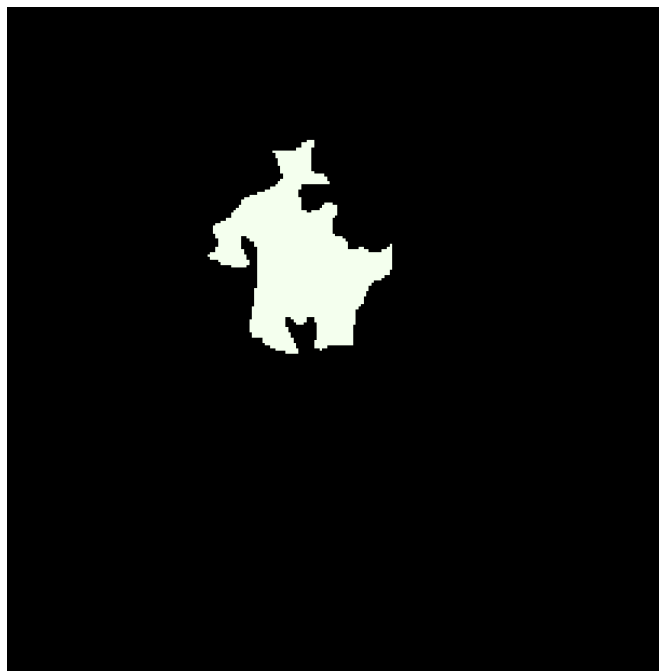
Žemo lygio gliomų segmentavimo tyrimui ir praktiniam realizavimui naudoti 110 pacientų (1 priedas) duomenys iš Vėžio nuotraukų archyvo (angl. *Cancer imaging archive*⁴) ir JAV Nacionalinio Vėžio Instituto (angl. *National Cancer Institute*⁵), paruošti pagal [MCC+17]. Duomenys pateikti iš penkių įstaigų: Henry'io Fordo ligoninės (45 pacientų), Thomaso Jeffersono Universiteto (16 pacientų), *Case Western* (14 pacientų), *Case Western* - šv. Josepho (34 pacientai) ir *UNC* (angl. *University of North Carolina*) (1 paciento). Siekiant geriau suprasti naviko vystymąsi buvo pasirinkti ikioperaciniai MRT vaizdai.

Tais atvejais, kai yra prieinami visi MRT modalumai, jie visi ir yra naudojami. Priešingu atveju naudojamas tik skysčio supresijos inversijos režimas (angl. *fluid-attenuated inversion recovery*) - *FLAIR* modalumas. *FLAIR* režimas gali būti naudojamas tiek su dvimačiais, tiek ir su trimačiais vaizdais. Be to, jis yra labai jautrus anomalijoms ligų diagnozavime: normalus smegenų skystis (angl. *cerebrospinal fluid, CSF*) nuotraukoje išlieka tamsus, o anomalijos – itin šviesios (9 pav. galime stebėti keturias ryškiai šviesios spalvos anomalijų sritis keturiuose *FLAIR* modalumo vaizduose). Tuo tarpu baltoji smegenų medžiaga *FLAIR* modalumo vaizduose atrodo tamsiai pilka, o galvos smegenų žievė – šviesiai pilka. Autoriai [BSM19] pastebėjo, jog iš visų analizuotų atvejų, 101 paciento duomenyse buvo galima rasti visus MRT modalumus. Segmentavimo modelio įvesties vaizdai saugomi *TIF* (angl. *Tagged Image File*⁶) formatu ir yra 256x256 dydžio. [MCC+17] apjungto duomenų rinkinio vaizdai buvo susegmentuoti rankiniu būdu (8 pav.) su *MATLAB*, su radiologijos sritimi susipažinusio ir medicininį išsilavinimą turinčio tyrėjo bei, to prireikus, ištaisyti ir patvirtinti skirtingų neuroradiologų. Rankinis *FLAIR* modalumo vaizdų žymėjimas vyko nubrėžiant *FLAIR* anomalijų ribas ant kiekvieno MRT pjūvio.

⁴ <https://www.cancerimagingarchive.net/>

⁵ <https://portal.gdc.cancer.gov/>

⁶ <https://www.adobe.io/open/standards/TIFF.html>



8 pav. Radiologų patvirtinta 256x256 dydžio⁷ segmentavimo kaukė.

3.2 Duomenų paruošimas

Straipsnio [BSM19] autoriai pastebėjo, jog, priklausomai nuo paciento, surinktuose duomenų rinkiniuose MRT vaizdai ženkliai skyrėsi savo dydžiu. Todėl MRT vaizdo sekos buvo paruoštos pašalinant kaukolės kontūrus (angl. *skull stripping*), kad visą dėmesį būtų galima skirti smegenų sritims. Tam buvo panaudotas giliojo mokymo pagrindu veikiantis algoritmas⁸, besiremiantis *FLAIR* modalumu ir geriausiai susegmentuoto vaizdo atveju pasiekiantis 0,9662 Dice'o koeficiento įvertį. Blogiausiu segmentavimo atveju šis įvertis buvo 0,9476. Taip pat duomenų paruošimo etape, remiantis vaizdo histograma, buvo normalizuoti audinių MR signalų intensyvumai. Galiausiai visiems duomenims buvo pritaikytas *Z-score* normalizavimas.

3.3 Įvertinimo kriterijai

Remiantis [MCC+17], vertinant modelį atsižvelgiama tik į *FLAIR* modalumo, o ne naviko sustiprinimo srities anomalijas, nes pastarosios ŽLG atvejuose pasitaiko retai. Šiuo pagrindu įvertinama naviko forma. Kaip įprasta galvos smegenų navikų segmentavimo uždaviniuose,

⁷ Ataskaitoje vaizdas išdidintas dėl didesnio aiškumo.

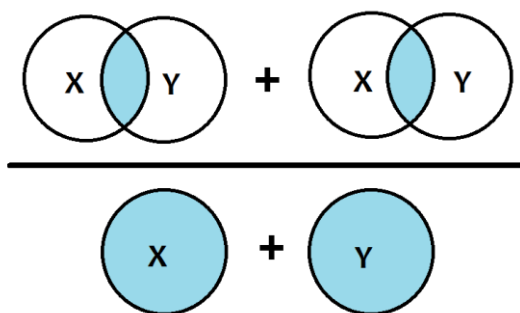
⁸ <https://github.com/MaciejMazurowski/brain-segmentation/tree/master/skull-stripping>

įvertinimui pasirinkta Dice'o koeficiento metrika (3.1), matuojanti, kokia modelio susegmentuotų sričių dalis persidengia su ekspertų sudarytomis segmentavimo kaukėmis.

Dice'o koeficiento formulė:

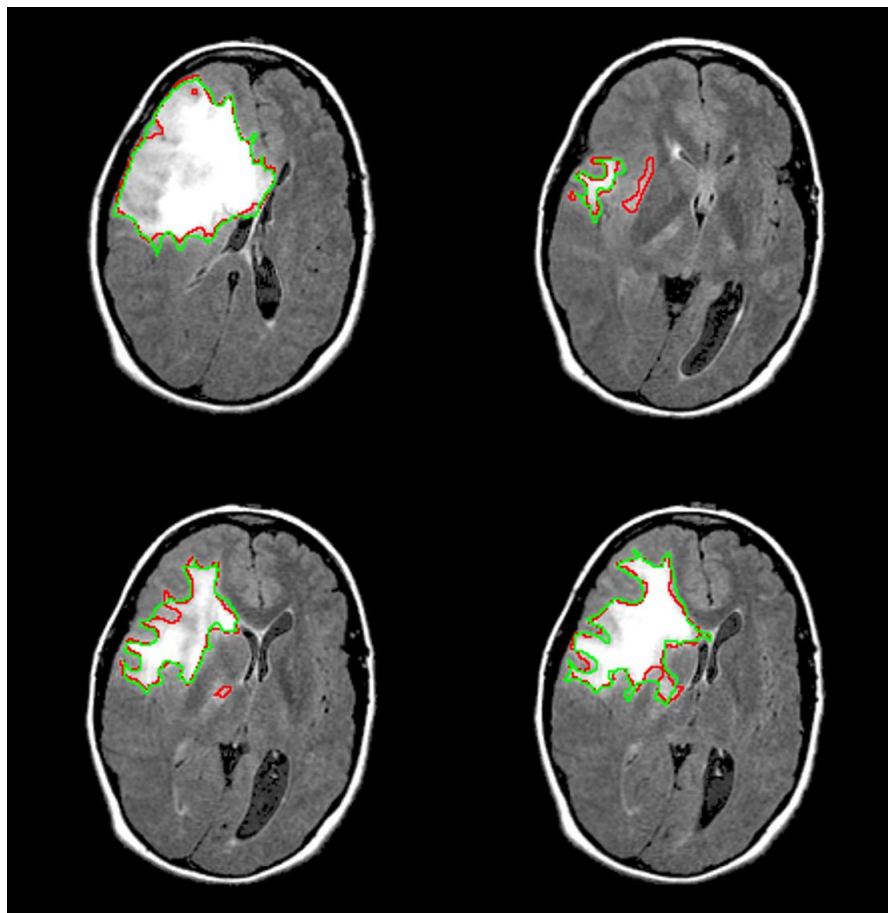
$$\frac{2 * |X \cap Y|}{|X| + |Y|} \quad (3.1)$$

X žymi modelio prognozuojamų anomalijos srities pikselių aibę, o Y – rankiniu būdu anomalijos sričiai priskirtų pikselių aibę (9 pav.). Skaitiklyje turime prognozuojamos pikselių aibės sankirtos su rankiniu būdu sužymėta pikselių aibe rezultata, padaugintą iš 2. Tuo tarpu vardiklyje skaičiuojama šių dviejų pikselių aibių suma. Nepersidengiančių aibių atveju skaitiklio, o tuo pačiu ir Dice'o įverčio reikšmė bus lygi 0.



9 pav. Į Dice'o įverčio skaičiavimą įtraukiamos aibės ir poaibiai pažymėti melsvai.

Žaliai susegmentuoti kontūrai (10 pav.) žymi radiologų rankiniu būdu sužymėtas sritis, o raudonos segmentavimo linijos žymi vieno iš modifikuotų modelių – Versija_2 (3 lentelė) – susegmentuotą sritį. Modelio ir jo modifikacijų įvertinimui buvo naudota kryžminė patikra, kuomet 110 pacientų duomenų rinkinys suskaidomas į 22 poabius po 5 pacientus. Kaip pastebi [BSM19], skirtingų radiologijos ekspertų viso ŽLG tipo naviko segmentavimo Dice'o koeficientas siekia 84 proc. su 2 proc. standartiniu nuokrypiu, tad šią ribą pasiekiantys modeliai reikšmingai demonstruoja automatinio segmentavimo galimybes.



10 pav. Žalios linijos žymi radiologų rankiniu būdu susegmentuotas sritis, o raudonos – modelio susegmentuotą sritį. [U-Net modelis su [15; 0,10] augmentavimo hiperparametru rinkiniu (3 lentelė)].

3.4 Modeliai

Eksperimentinės tyrimų dalies modeliai įgyvendinti su *Python 3.7.7* ir *PyTorch 1.3.1* biblioteka. Tinklai apmokyti naudojant *Lenovo ThinkPad P70* su dvejais GPU (*GPU 0: Intel® HD Graphics P530*, *GPU 1: NVIDIA Quadro M3000M*). Nelygiagrečių skaičiavimų operacijoms naudotas CPU – *Intel® Xeon® E3-1505M v5*. Operacinė sistema – *Windows 10 Pro 64-bit*, operatyvioji atmintis – *64,0 GB*. Dalis intensyvesnių (GPU atminties resursų sunaudojimo prasme) skaičiavimų atlikta su *Tesla* architektūros *P100 GPU* *Kaggle* platformoje. Galiausiai duomenų vizualizavimui naudota *Matplotlib* biblioteka.

Iš viso žemo lygio gliomų segmentavimui realizuotos 4 skirtingos architektūros: *U-Net*, *SegNet*, *SegNet* su rekurentiniais liekanų blokais bei *U-Net* su dėmesio vartais (angl. *attention U-Net*). Pradinių modelio filtrų skaičius nurodytas modelio hiperparametru lentelių pavadinimuose.

Geriausių rezultatų vidutinio validavimo Dice'o įverčio prasme pasiekusios modelio modifikacijos pažymėtos žalia spalva.

Siekiant rasti geriausius KNT hiperparametrus, pasirinktas *grid search* tipo hiperparametrų paieškos metodas, kuomet testuojamos to paties hiperparametro skirtingos reikšmės yra panašiais intervalais nutolusios viena nuo kitos. Visgi, turint daugiau resursų ir tiriant didesnę hiperparametrų bei galimų jų reikšmių skaičių, atsitiktinės paieškos (angl. *random search*) metodas turėtų duoti geresnius rezultatus.

Tiriami modeliai apmokyti su 3 hiperparametrų reikšmių kombinacijomis:

- Duomenų rinkiniai (angl. *mini-batch size*): [8; 16; 32]
- Epochos: [25; 50; 75]
- Mokymo greičiai: [0,0001; 0,001]

Duomenų rinkinys nurodo, kiek apmokymo aibės vienetų prasklis pro mokomą tinklą vienos iteracijos metu. Didesnė reikšmė pareikalaus daugiau atminties resursų. Šio hiperparametro tiriamomis reikšmėmis pasirinkti skaičiai su laipsnio pagrindu 2 ir sveikuoju rodikliu, nes taip efektyviau išnaudojami *CPU* ir *GPU* resursai. Pabandžius modelius apmokyti su duomenų reikšmėmis [10; 30] pastebėta, kad daugumoje atvejų geresni rezultatai pasiekiami šias reikšmes pakeitus į [8; 32]. Dėl ribotų atminties resursų buvo koncentruojamasi į modelius, kurių pradinių filtrų skaičiaus reikšmė neviršytų 32.

3.4.1 *U-Net* tyrimas ir rezultatai

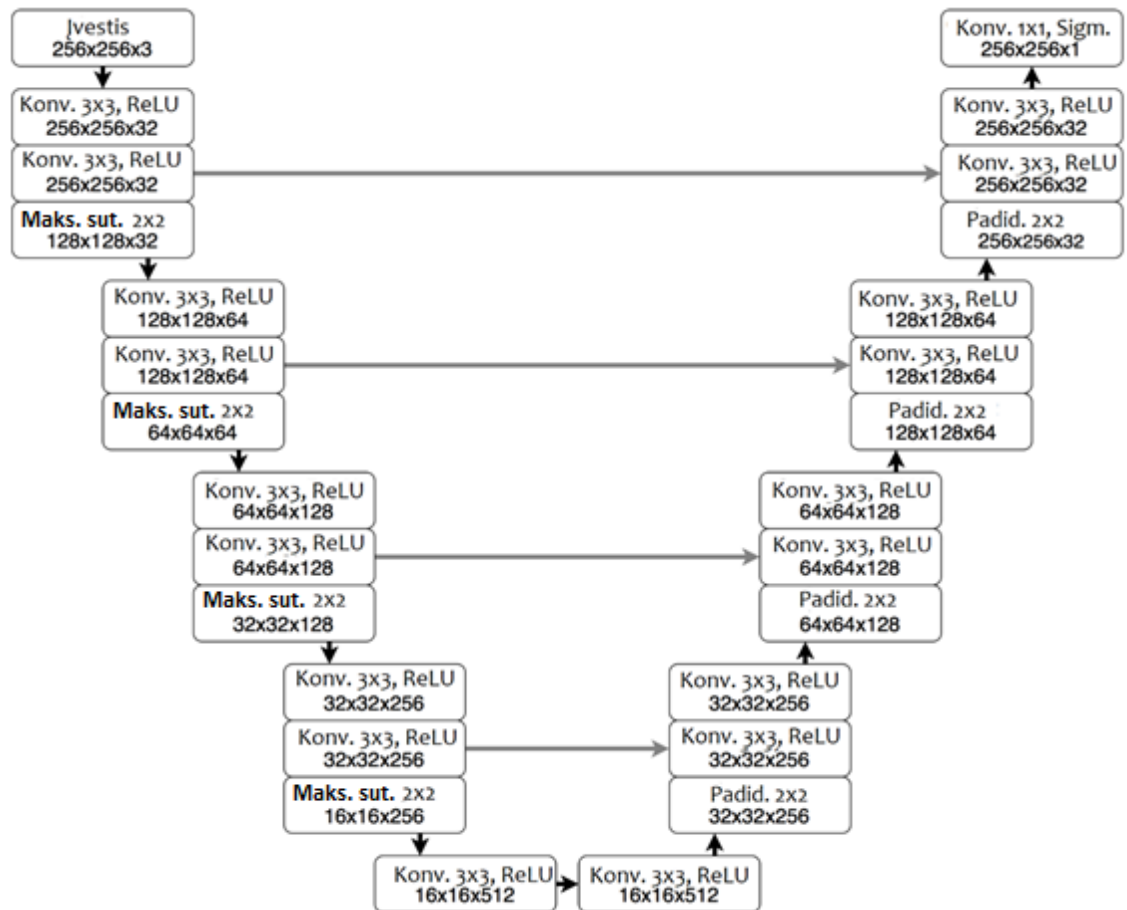
Šiame darbe realizuoti *U-Net* tipo architektūros [BSM19] modeliai, sudaryti iš 4 enkoderio ir 4 dekoderio blokų, ir viduriniojo, t.y. enkoderio ir dekoderio dalis apjungiančio, bloko. Stačiakampiuose tinklo blokuose (11 pav.) po sluoksnių pavadinimais sužymėti jų išeities matmenys. Siekiant ištirti skirtingų su augmentavimu susijusių hiperparametrų įtaką modelių įverčiams, be pradinės Mateuszo Buda'os modelio versijos^{9 10 11} apmokytos dar 17 modelio versijų (4 lentelė) su modifikuotais hiperparametrais. Pradinė MRT nuotrauka, prieš patekdama į enkoderio dalies įvestį, yra sumažinama nuo 256x256 iki 224x224 dydžio. Tokiu būdu dalinai sutaupomi atminties resursai. Tuomet palaipsniui taikoma dvimatė konvoliucijos operacija su 3x3 dydžio filtrais, *ReLU* aktyvacijos funkcija ir dvimatis maksimalių reikšmių sutelkimo sluoksnis.

⁹ <https://github.com/mateuszbeda/brain-segmentation-pytorch>

¹⁰ <https://github.com/MaciejMazurowski/brain-segmentation>

¹¹ <https://www.kaggle.com/mateuszbeda/brain-segmentation-pytorch>

Enkoderio dalies maksimalių reikšmių sutelkimo ir dekoderio dalies išdidinimo sluoksniuose poslinkių reikšmės nustatytos kaip 2. Kaip nuostolių funkcija taikyta Dice'o nuostolių funkcija. Sumažinę MRT vaizdo raišką enkoderio dalyje, dekoderio dalies pabaigoje susigrąžiname pradinį 224x224 jau susegmentuoto vaizdo dydį, išeityje palikdami tik vieną kanalą – *FLAIR* modalumą.



11 pav. *U-Net* tipo architektūra su ŽLG susijusių anomalijų segmentavimui.

Šaltinis: [<https://github.com/MaciejMazurowski/brain-segmentation>].

Pradiniame *U-Net* modelyje (3 ir 4 lentelėje paryškintame mėlsva spalva) su 50 epochų, 16 duomenų rinkiniu ir 0,0001 mokymo greičiu iširta su duomenų augmentavimu susijusių hiperparametrų įtaka Dice'o įverčiui.

- Augmentavimo dydis: [0,05; 0,1; 0,15; 0,20]
- Pasukimo kampas: [5; 10; 15; 20]

3 lentelė. *U-Net*, 32 pradiniai filtrai, su duomenų augmentavimu susiję hiperparametrai.

Pasukimo kampas	Augmentavimo dydis	Geriausias valid. Dice'o įvertis
5	0,05	0,8891
5	0,10	0,8747
5	0,15	0,8722
5	0,20	0,8831
10	0,05	0,8658
10	0,10	0,8835
10	0,15	0,8775
10	0,20	0,8912
15	0,05	0,8815
15	0,10	0,8788
15	0,15	0,8768
15	0,20	0,8825
20	0,05	0,8774
20	0,10	0,9018
20	0,15	0,8717
20	0,20	0,8853

Dalį modelio versijų pamėginus apmokyti antrą sykį, matyti, jog pasirinktų augmentavimo hiperparametrų rinkiniai gali įtakoti Dice'o koeficiento įvertį, jį pagerindami arba supratindami apytiksliai 1 proc. Modelio versija su 20 laipsnių pasukimo kampu ir 0,10 augmentavimo dydžiu pasiekė geriausių rezultatų Dice'o įverčio atžvilgiu, tačiau, lyginant su kitomis versijomis, modelio validavimo nuostolių įverčio reikšmė buvo 0,05-0,10 didesnė, t.y. prastesnė. Dėl šios priežasties gali būti prasminga pasirinkti mažesnes duomenų augmentavimo hiperparametrų reikšmes, kaip tą daro [BSM19] autoriai. Tokių pasukimo kampo ir augmentavimo dydžio hiperparametrų rinkinių pavyzdžiai yra [10; 0,20] arba [15; 0,05]. Norėdami ištirti augmentavimo hiperparametrų reikšmes kartu su trimis anksčiau pasirinktais hiperparametrais (duomenų rinkiniu, epochomis bei mokymo greičiu), turėtume patikrinti 288 skirtingas kombinacijas kiekvienai mūsų tiriamai architektūrai, kas nėra efektyvu resursų naudojimo atžvilgiu, todėl toliau nagrinėjamuose modeliuose naudosime [15; 0,05] augmentavimo hiperparametrų rinkinį.

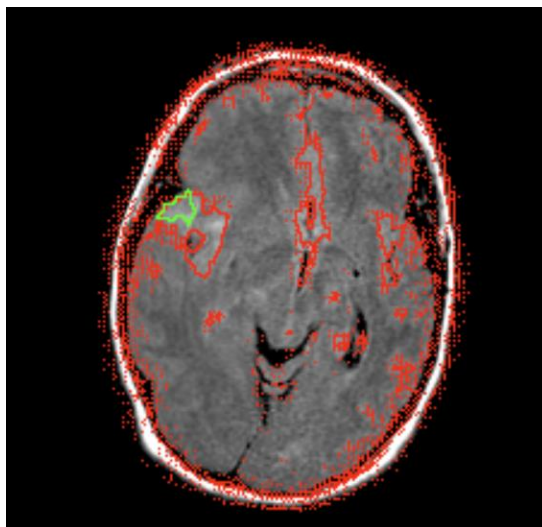
Ištirus 32 pradinių filtrų *U-Net* architektūrą su visomis pasirinktų hiperparametrų kombinacijomis matyti, jog geriausi Dice'o įverčio rezultatai pasiekiami 7 ir 16 modeliuose (4 lentelė): 0,8902 ir 0,8900. Modeliai su 25 epochų hiperparametro reikšme parodė prastesnius validavimo nuostolių rezultatus, kuriuos, padidinus apmokomo modelio epochų skaičių, būtų galima pagerinti reikšmingai nesuprastinant Dice'o įverčio reikšmės. Geriausius validavimo nuostolių rezultatus pasiekė modeliai su 0,001 mokymo greičiu. Todėl renkantis tinklą reikėtų atsižvelgti ir į modelių versijas 12 ir 14, kurių validavimo nuostoliai siekė tik 0,10 ir 0,15, o jų Dice'o įverčiai nedaug tesiskyrė nuo geriausių įverčių pasiekusių modelių. Silpniausiai pagal Dice'o įvertį pasirodžiusio *U-Net* modelio rezultatas pasiekė 0,8475.

4 lentelė. *U-Net*, 32 pradiniai filtrai.

Modelis	Epochos	Duomenų rinkinys	Mokymo greitis	Geriausias valid. Dice'o įvertis
1	25	8	0,0001	0,8883
2	25	8	0,001	0,8812
3	25	16	0,0001	0,8875
4	25	16	0,001	0,8475
5	25	32	0,0001	0,8671
6	25	32	0,001	0,8642
7	50	8	0,0001	0,8902
8	50	8	0,001	0,8660
9	50	16	0,0001	0,8815
10	50	16	0,001	0,8685
11	50	32	0,0001	0,8662
12	50	32	0,001	0,8871
13	75	8	0,0001	0,8894
14	75	8	0,001	0,8897
15	75	16	0,0001	0,8871
16	75	16	0,001	0,8900
17	75	32	0,0001	0,8717
18	75	32	0,001	0,8866

Visos 18 modelio versijų buvo apmokytos su *Adam* optimizavimo algoritmu. *Adam* taikymas leidžia pasiekti geresnius rezultatus lyginant su *SGD* (angl. *stochastic gradient descent*) optimizavimu, kuris pasirodė netinkamas šiam modeliui. Pabandžius pritaikyti *SGD*, gauti prasčiausi rezultatai lyginant su 4 lentelėje gautais Dice'o įverčiais taikant *Adam*: vidutinis Dice'o koeficiento įvertis – tik 0,1164 (tai geriausiai pasimato (13 pav.), kur net plika akimi matyti, kad

klaidingai klasifikuotų pikselių aibė bus palyginti didelė), Dice'o nuostolių funkcijos įvertis - 0,9014, o validavimo nuostolių - 0,9590.



13 pav. Nuotrauka susegmentuota taikant *SGD* optimizavimą.

3.4.2 *U-Net* su dėmesio vartais tyrimai ir rezultatai

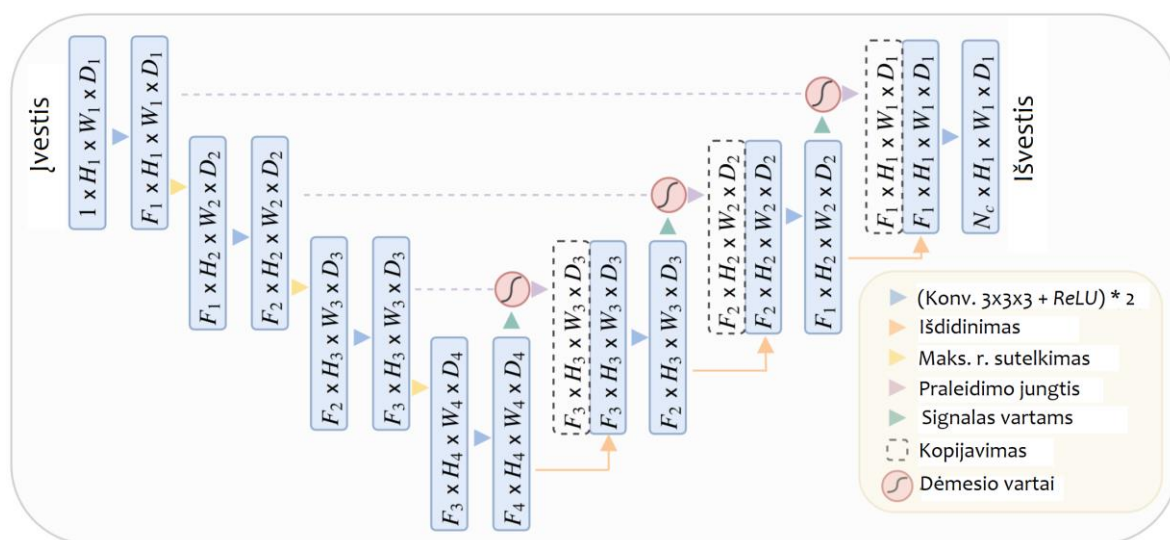
Modeliai apmokyti remiantis [OSF+18] pradinės architektūros modeliais^{12 13}. Nors pradinis architektūros siūlymas yra trimatis ir susidedantis iš 3 enkoderio ir 3 dekoderio blokų, šią architektūrą galima naudoti ir su dvimačių smegenų pjūvių nuotraukomis. Architektūros brėžinyje (14 pav.) rausvais apskritimais pažymėti dėmesio vartų moduliai. Į jų vidų patenka atitinkama praleidimo jungtis iš to paties *U-Net* gylio sluoksnio ir išdidintas bei konvoliucijos operaciją praėjęs signalas vartams, siunčiamas iš gilesnio (nei adresuojamų dėmesio vartų) sluoksnio. Dėmesio vartų modulyje, prieš tai kiekvienam iš jų atskirai pritaikius konvoliucijos ir normalizavimo operacijas, šie dvejų įvesties kanalų rezultatai yra sudedami ir praleidžiami pro *ReLU* funkciją. Tuomet gautam rezultatui ir vėl pritaikomos konvoliucijos bei normalizavimo operacijos. Galiausiai rezultatas tampa įvestimi *sigmoid* funkcijoje, kurios rezultatas sudauginamas su pradine, į dėmesio vartus patekusios, praleidimo jungties reikšme. Būtent šis, kiekvieno (išskyrus žemiausiąjį) *U-Net* gylio sluoksnyje, pasikartojantis dėmesio vartų modulis su

¹² <https://github.com/ozan-oktay/Attention-Gated-Networks>

¹³ https://github.com/LeeJunHyun/Image_Segmentation

į jį įeinančiomis reikšmėmis ir yra vienintelis ankstesniame skirsnyje nagrinėtos klasikinės *U-Net* architektūros papildymas.

Realizuotuose ir eksperimentiškai tirtuose modeliuose įkomponuoti 4 dėmesio vartų moduliai, pro kuriuos praeina *U-Net* praleidimo jungtys. Dėmesio vartų modulių taikymas neapsiriboja *U-Net* tinklu ir gali būti praplėstas kitose architektūrose – pavyzdžiui, žinant, jog dėmesio vartus galime naudoti su *VGG-16* architektūra, kurios pirmieji konvoliuciniai sluoksniai sudaro *SegNet* enkoderį, galime spręsti, kad juos pritaikyti būtų galima ir *SegNet* tinklui.



14 pav. 3 enkoderio ir 3 dekoderio blokų trimatis *U-Net* su dėmesio vartais. Šaltinis: [OSF+18].

Dėl ribotų atminties resursų, *U-Net* su dėmesio vartais versijos, naudojančios 32 duomenų rinkinius, buvo apmokytos *Kaggle* platformoje. Geriausias modelio pasiektas rezultatas – 0,9008 Dice'o įvertis su [50; 8; 0,0001] hiperparametrų rinkiniu. Žemesnės validavimo nuostolių reikšmės ir vėl buvo pastebėtos tuose modeliuose, kurių mokymo greičio hiperparametro reikšmė buvo 0,001 (tokios modelių versijos (5 lentelėje) pažymėtos melsvai).

5 lentelė. *U-Net* su dėmesio vartais, 32 pradiniai filtrai.

Modelis	Epochos	Duomenų rinkinys	Mokymo greitis	Geriausias valid. Dice'o įvertis
1	25	8	0,0001	0,8849
2	25	8	0,001	0,8727
3	25	16	0,0001	0,8724
4	25	16	0,001	0,8588
5	25	32	0,0001	0,8684
6	25	32	0,001	0,8873
7	50	8	0,0001	0,9008
8	50	8	0,001	0,8805
9	50	16	0,0001	0,8830
10	50	16	0,001	0,8716
11	50	32	0,0001	0,8722
12	50	32	0,001	0,8771
13	75	8	0,0001	0,8900
14	75	8	0,001	0,8782
15	75	16	0,0001	0,8779
16	75	16	0,001	0,8838
17	75	32	0,0001	0,8770
18	75	32	0,001	0,8856

3.4.3 *SegNet* tyrimas ir rezultatai

Tinklo modeliai realizuoti pagal pirminę tinklo struktūrą [BKC16], paskutinį sluoksnį iš *softmax* pakeičiant į *sigmoid*. Pradiniame tinklo sluoksnyje naudojami 8 filtrai, o jų skaičius padvigubėja po kiekvieno *SegNet* enkoderio dalies bloko, kurį pradinėje dalyje sudaro 2 konvoliucijos operacijos su po jų einančiu sutelkimo sluoksniu arba, vėlesnėje dalyje, 3 konvoliucijos operacijos su tuo pačiu sutelkimo sluoksniu. Skirtingai nei *U-Net* realizacijos atveju, naudodami *Pytorch* maksimalaus reikšmių sutelkimo funkciją *MaxPool2d*, nurodome papildomą parametą *return_indices* ir jo reikšmę nustatome kaip *True*. Tokiu būdu gausime prieigą prie maksimalių reikšmių indeksų bei jų dydžių, kurie reikalingi išskleidžiant maksimalias reikšmes dekoderio dalyje.

Aukščiausias Dice'o panašumo koeficiento įvertis pasiektas 16 modelio versijoje – 0,8261, apmokant modelį 75 epochas po 16 nuotraukų per iteraciją su 0,001 mokymo greičiu. Šio *SegNet* modelio rezultatas pralenkia [CLY+18] pasiektą 0,824 viso naviko segmentavimo Dice'o koeficientą. Eksperimentuojant su optimizavimo algoritmo keitimu iš *Adam* į *AdamW* nepastebėta reikšminga įtaka rezultatams, todėl visos *SegNet* modelių versijos (6 lentelė) apmokytos taikant

Adam. Prasčiausias Dice'o koeficiento įvertis pasiektas 17 modelio versijoje – 0,6021, taikant [75; 32; 0,0001] hiperparametrų rinkinį.

6 lentelė. SegNet, 8 pradiniai filtrai.

Modelis	Epochos	Duomenų rinkinys	Mokymo greitis	Geriausias valid. Dice'o įvertis
1	25	8	0,0001	0,7310
2	25	8	0,001	0,7155
3	25	16	0,0001	0,6550
4	25	16	0,001	0,7614
5	25	32	0,0001	0,6541
6	25	32	0,001	0,7675
7	50	8	0,0001	0,7965
8	50	8	0,001	0,7801
9	50	16	0,0001	0,7617
10	50	16	0,001	0,7725
11	50	32	0,0001	0,6678
12	50	32	0,001	0,7934
13	75	8	0,0001	0,8235
14	75	8	0,001	0,7898
15	75	16	0,0001	0,7745
16	75	16	0,001	0,8261
17	75	32	0,0001	0,6021
18	75	32	0,001	0,7763

3.4.4 SegNet su rekurentiniais liekanų blokais tyrimas ir rezultatai

Ši SegNet tinklo modifikacija sukurta 2 pradines SegNet konvoliucijos operacijas pakeitus rekurentiniu liekanų bloku¹⁴ (angl. *recurrent residual block*), įkvėptu R2U-Net tinklo [AHY+18]. Taigi pakeista tik pati enkoderio dalies pradžia (t.y. pirmasis aukščiau aptarto SegNet enkoderio blokas) ir dekoderio dalies pabaiga (t.y. paskutinis dekoderio blokas). Tačiau pradiniai rezultatai rodo Dice'o įverčio pagerėjimą lyginant su paprastomis SegNet architektūros versijomis. Geriausias Dice'o koeficiento rezultatas gautas 75 epochų modelyje – 0,8521, kai tuo tarpu geriausias SegNet Dice'o įvertis buvo 0,8261.

¹⁴ https://github.com/LeeJunHyun/Image_Segmentation

Ekspserimentuojant su *AdamW* optimizavimo algoritmu pastebėti ženkliai prastesni rezultatai lyginant su *Adam*. Geriausi validavimo nuostoliai pasiekti naudojant duomenų rinkinį su 32 nuotraukomis – geriausią Dice'o įvertį pasiekusioje versijoje ši reikšmė siekia 0,1443.

Geriausius validavimo nuostolių rezultatus pasiekė modeliai 12, 17 ir 18, kuriuose geriausios validavimo nuostolių reikšmės svyravo nuo 0,14 iki 0,23. Visgi dalis kitų modelių, su mažesnėmis epochų skaičiaus hiperparametro reikšmėmis, demonstravo suprastėjusius validavimo nuostolius (pav. 15), lyginant su *U-Net* modeliu ir jo modifikacija.

```

epocha 25 | nuostoliai: 0.05387402840861232
epocha 25 | validavimo nuostoliai: 0.5788106081031618
epocha 25 | validavimo Dice įvertis: 0.5241305009291192

Geriausias validavimo vidutinis Dice'o įvertis: 0.750423

```

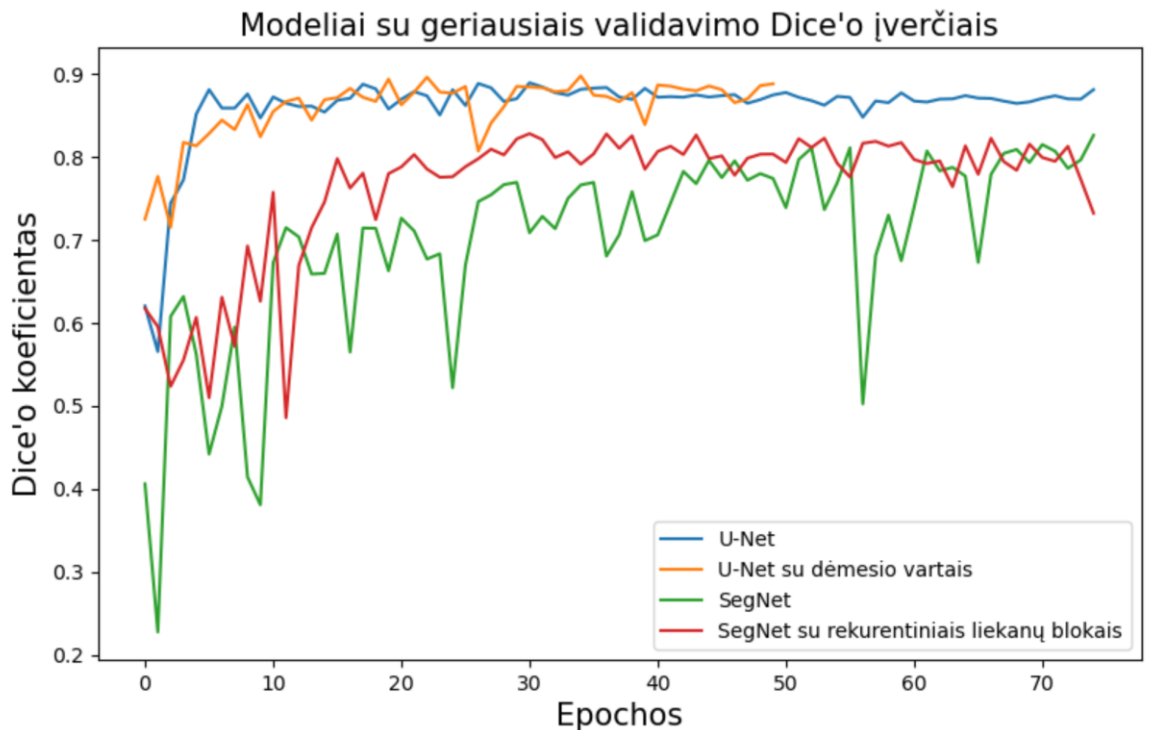
15 pav. Prastesni modelių validavimo nuostolių rezultatai.

7 lentelė. SegNet su rekurentiniais liekanų blokais, 8 pradiniai filtrai.

Modelis	Epochos	Duomenų rinkinys	Mokymo greitis	Geriausias valid. Dice'o įvertis
1	25	8	0,0001	0,4929
2	25	8	0,001	0,8202
3	25	16	0,0001	0,7551
4	25	16	0,001	0,7426
5	25	32	0,0001	0,6785
6	25	32	0,001	0,7765
7	50	8	0,0001	0,7024
8	50	8	0,001	0,6760
9	50	16	0,0001	0,7820
10	50	16	0,001	0,7895
11	50	32	0,0001	0,6831
12	50	32	0,001	0,7317
13	75	8	0,0001	0,8281
14	75	8	0,001	0,8487
15	75	16	0,0001	0,8403
16	75	16	0,001	0,7898
17	75	32	0,0001	0,7326
18	75	32	0,001	0,8521

3.5 Rezultatų įvertinimas ir palyginimas

Geriausią Dice'o panašumo koeficiento rezultatą pasiekė vienas iš *U-Net* su dėmesio vartais modelių (16 pav.). Pradinės architektūros *SegNet* modeliai demonstravo prasčiausius rezultatus, tačiau tinklo modifikacija pasitelkiant rekurentinius liekanų blokus pagerino *SegNet* segmentavimo rezultatus. Visose architektūrose geriausi rezultatai pasiekti naudojant *Adam* optimizavimo algoritmą. Greičiausiai apmokoma viena *SegNet* tinklo epocha, o lėčiausiai – *U-Net* su dėmesio vartais tinklo epocha.



16 pav. Modeliai su geriausius rezultatus davusiais hiperparametrų rinkiniais.

Kaip ir tikėtasi, *SegNet* architektūra bei jos modifikacija pareikalauja mažiausio parametru skaičiaus (8 lentelė). *SegNet* modelį lyginant su daugiausiai parametru turinčiu tinklu – *U-Net* su dėmesio vartais, pastebime beveik 9 kartais mažesnę tinklo parametru skaičiaus reikšmę. Tai viena priežasčių, kodėl *U-Net* su dėmesio vartais pareikalauja daugiausiai resursu, tačiau pasiekia geriausią Dice'o įvertį.

8 lentelė. Parametrų skaičiaus palyginimas skirtingose architektūrose.

Architektūra	Parametrai
<i>U-Net</i>	7763041
<i>U-Net</i> su dėmesio vartais	8726077
<i>SegNet</i>	981515
<i>SegNet</i> su rekurentiniais liekanų blokais	986708

4. Išvados

Darbo eigoje išanalizuotos pažangiausios KNT architektūros, tinkamos smegenų navikų segmentavimo užduočiai bei eksperimentiškai ištirtas ŽLG semantinio segmentavimo uždavinys. Pasiūlyta galima tinklo architektūros modifikacija pradiniam *SegNet* tinklui, jo enkoderio ir dekoderio dalyse pridėdant rekurentinius liekanų blokus, taip pagerinant *SegNet* Dice'o panašumo koeficiento įvertį. Praktinis *U-Net* realizavimas ir literatūros analizė parodė šios architektūros tipo tinklų tinkamumą gliomų segmentavimo uždaviniams spręsti. Panaudojant *U-Net* su dėmesio vartais, pradinio *U-Net* modelio rezultatas pagerintas nuo 0,8815 iki 0,9034 Dice'o panašumo koeficiento įverčio.

Iš gautų rezultatų galime daryti tokias išvadas:

1. Reikšmingiausi segmentavimo sprendimai yra įtakoti pilnai konvoliucinio tinklo ir remiasi enkoderio-dekoderio architektūra.
2. *U-Net* su dėmesio vartais apmokymas užtrunka ilgiausiai bei reikalauja daugiausiai atminties resursų, ypač lyginant su *SegNet*.
3. Dėmesio vartų modulių integravimas į *U-Net* leidžia pasiekti geriausią Dice'o įvertį lyginant su kitų ištirtų architektūrų modeliais.
4. *SegNet* segmentavimo rezultatai gali būti pagerinti pridėdant pasiūlytą rekurentinių liekanų blokų modifikaciją.

Didesnis MRT nuotraukų kiekis, didesni skaičiavimo resursai ir jų tiesiogiai įtakojama galimybė hiperparametrų radimui taikyti atsitiktinę paiešką leistų dar labiau pagerinti šiuo metu geriausią pasiektą Dice'o koeficiento įvertį.

Literatūra:

- [AGH+17] Akkus Z., Galimzianova A., Hoogi A., Rubin D. L., Erickson B. J. Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. *Journal of Digital Imaging*. Vol. 30-4. 2017.
- [AHY+18] Alom M. Z., Hasan M., Yakopcic C., Taha T. M., Asari V. K. Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation. 2018. arXiv:1802.06955v5 [cs.CV]
- [BKC16] Badrinarayanan V., Kendall A., Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. 2016. arXiv:1511.00561v3 [cs.CV]
- [BMS18] Banerjee S., Mitra S., Shankar B. U. Multi-Planar Spatial-ConvNet for Segmentation and Survival Prediction in Brain Cancer. *Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge*. 2018. P. 36-45.
- [BPF+18] Benson E., Pound M. P., French A. P., Jackson A. S., Pridmore T. P. Deep Hourglass for Brain Tumor Segmentation. *Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge*. 2018. P. 46-53.
- [BSM19] Buda M., Saha A., Mazurowski M. A. Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm. 2019. arXiv: 1906.03720v1 [eess.IV]
- [CVG+18] Cabezas M., Valverde S., Gonzalez-Villa S., Cleriques A., Salem M., Kushibar K., Bernal J., Oliver A., Salvi J., Llado X. Survival prediction using ensemble tumor segmentation and transfer learning. *Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge*. 2018. P. 54-62.
- [CVJ+18] Choudhury A. R., Vanguri R., Jambawalikar S. R., Kumar P. Segmentation of Brain Tumors using DeepLabv3+. *Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge*. 2018. P. 111-118.
- [CPA+16] Casamitjana A., Puch S., Aduriz A., Sayrol E., Vilaplana V. 3D Convolutional Networks for Brain Tumor Segmentation.

<https://imatge.upc.edu/web/sites/default/files/pub/cCasamitjana16.pdf>

- [CCL+18] Caver E., Chang, Liu, Zong W., Dai Z., Wen N. Automatic Brain Tumor Segmentation Using a U-net Neural Network. *Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge*. 2018. P. 63-73.
- [CVF+18] Chandra S., Vakalopoulou M., Fidon L., Battistella E., Estienne T., Sun R., Robert C., Deutch E., Paragios N. Context Aware 3-D Residual Networks for Brain Tumor Segmentation. *Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge*. 2018. P. 74-82.
- [CLY+18] Chang Y., Lin Z., Yang T., Huang T. Automatic segmentation of brain tumor from 3D MR images using a 2D convolutional neural network. *Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge*. 2018. P. 83-90.
- [CYA+19] Cardenas C. E., Yang J., Anderson B. M., Court L. E., Brock K. B. Advances in Auto-Segmentation. *Seminars in Radiation Oncology*. No. 29. 2019. P. 185-197.
- [Fri18] Fridman N. Brain Tumor Detection and Segmentation Using Deep Learning U-Net on Multi Modal MRI. *Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge*. 2018. P. 135-143.
- [GJ12] Goodenberger M., Jenkins R. B. Genetics of adult glioma. *Cancer Genetics*. 2012. Nr. 12. P. 613-621.
- [HZR+15] He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. 2015. arXiv:1512.03385v1 [cs.CV]
- [IKW+18] Isensee F., Kickingereder P., Wick W., Bendszus M., Maier-Hein K. H., No New-Net. *Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge*. 2018. P. 222-231.
- [KBC16] Kendall A., Badrinarayanan V., Cipolla R. Bayesian SegNet: Model Uncertainty in Deep Convolutional Encoder-Decoder Architectures for Scene Understanding. 2016. arXiv:1511.02680v2 [cs.CV]
- [KLN+17] Kamnitsas K., Ledig C., Newcombe V. F. J., Simpson J. P., Kane A. D., Menon D. K., Rueckert D., Glocker B. Efficient Multi-Scale 3D CNN with

- fully connected CRF for Accurate Brain Lesion Segmentation. 2017. arXiv:1603.05959v3 [cs.CV]
- [KSH12] Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*. 2012. P. 1097-1105.
- [LL19] Lundervold A. S., Lundervold A. An overview of deep learning in medical imaging focusing on MRI. *Z Med Phys* 29. 2019. P. 102-127.
- [LPL+18] Liu J., Pan Y., Li M., Chen Z, Tang L., Lu C., Wang J. Applications of Deep Learning to MRI Images: A Survey. *Big Data Mining And Analytics*. Vol. 1-1. 2018.
- [LWY+17] Li Z., Wang Y., Yu J., Shi Z., Guo Y., Chen L., Mao Y. Low-Grade Glioma Segmentation Based on CNN with Fully Connected CRF. *Journal of Healthcare Engineering*. 2017.
- [MCC+17] Mazurowski M. A., Clark K., Czarnek N. M., Shamsesfandabadi P., Peters K. B., Saha A. Radiogenomics of lower-grade glioma: algorithmically-assessed tumor shape is associated with tumor genomic subtypes and patient outcomes in a multi-institutional study with The Cancer Genome Atlas data. *Journal of Neuro-Oncology*. 2017.
- [MJB+15] Menze B. H., Jakab A., Bauer S. et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging*. Vol. 34. No. 10. 2015. P. 1993-2024.
- [Myr18] Myronenko A., 3D MRI brain tumor segmentation using autoencoder regularization. *Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge*. 2018. P. 349-356.
- [OSF+18] Oktay O., Schempler J., Folgoc L., Lee M., Heinrich M., Misawa K., Mori K., McDonagh S., Y Hammerla N., Kainz B., Glocker B., Rueckert D. Attention U-Net: Learning Where to Look for the Pancreas. 2018, arXiv:1804.03999v3 [cs.CV]
- [RFB15] Ronneberger O., Fischer P., Brox T.: U-Net: convolutional networks for biomedical image segmentation. 2015, arXiv:1505.04597v1 [cs.CV]

- [SA17] Shen L., Anderson T. Multimodal Brain MRI Tumor Segmentation via Convolutional Neural Networks.
<http://cs231n.stanford.edu/reports/2017/pdfs/512.pdf>
- [SZ15] Simonyan, K., Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. 2015. arXiv:1409.1556v6 [cs.CV]
- [ZCD+18] Zhou C., Chen S., Ding C., Tao D. Learning Contextual and Attentive Information for Brain Tumor Segmentation. *Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge*. 2018. P. 571-578.

1 priedas.

110 pacientų, kurių duomenys buvo naudoti tyrime, sąrašas pagal identifikacijos numerį¹⁵
[MCC+17]:

TCGA-CS-4941, TCGA-CS-4942, TCGA-CS-4943, TCGA-CS-4944, TCGA-CS-5393,
TCGA-CS-5395, TCGA-CS-5396, TCGA-CS-5397, TCGA-CS-6186, TCGA-CS-6188,
TCGA-CS-6290, TCGA-CS-6665, TCGA-CS-6666, TCGA-CS-6667, TCGA-CS-6668,
TCGA-CS-6669, TCGA-DU-5849, TCGA-DU-5851, TCGA-DU-5852, TCGA-DU-5853,
TCGA-DU-5854, TCGA-DU-5855, TCGA-DU-5871, TCGA-DU-5872, TCGA-DU-5874,
TCGA-DU-6399, TCGA-DU-6400, TCGA-DU-6401, TCGA-DU-6404, TCGA-DU-6405,
TCGA-DU-6407, TCGA-DU-6408, TCGA-DU-7008, TCGA-DU-7010, TCGA-DU-7013,
TCGA-DU-7014, TCGA-DU-7018, TCGA-DU-7019, TCGA-DU-7294, TCGA-DU-7298,
TCGA-DU-7299, TCGA-DU-7300, TCGA-DU-7301, TCGA-DU-7302, TCGA-DU-7304,
TCGA-DU-7306, TCGA-DU-7309, TCGA-DU-8162, TCGA-DU-8163, TCGA-DU-8164,
TCGA-DU-8165, TCGA-DU-8166, TCGA-DU-8167, TCGA-DU-8168, TCGA-DU-A5TP,
TCGA-DU-A5TR, TCGA-DU-A5TS, TCGA-DU-A5TT, TCGA-DU-A5TU, TCGA-DU-A5TW,
TCGA-DU-A5TY, TCGA-EZ-7264, TCGA-FG-5962, TCGA-FG-5964, TCGA-FG-6688,
TCGA-FG-6689, TCGA-FG-6690, TCGA-FG-6691, TCGA-FG-6692, TCGA-FG-7634,
TCGA-FG-7637, TCGA-FG-7643, TCGA-FG-8189, TCGA-FG-A4MT, TCGA-FG-A4MU,
TCGA-FG-A60K, TCGA-HT-7473, TCGA-HT-7475, TCGA-HT-7602, TCGA-HT-7605,
TCGA-HT-7608, TCGA-HT-7616, TCGA-HT-7680, TCGA-HT-7684, TCGA-HT-7686,
TCGA-HT-7690, TCGA-HT-7692, TCGA-HT-7693, TCGA-HT-7694, TCGA-HT-7855,
TCGA-HT-7856, TCGA-HT-7860, TCGA-HT-7874, TCGA-HT-7877, TCGA-HT-7879,
TCGA-HT-7881, TCGA-HT-7882, TCGA-HT-7884, TCGA-HT-8018, TCGA-HT-8105,
TCGA-HT-8106, TCGA-HT-8107, TCGA-HT-8111, TCGA-HT-8113, TCGA-HT-8114,
TCGA-HT-8563, TCGA-HT-A5RC, TCGA-HT-A616, TCGA-HT-A61A, TCGA-HT-A61B

¹⁵ <https://portal.gdc.cancer.gov/>